

Proceedings of the Twenty-Sixth Water Reactor Safety Information Meeting

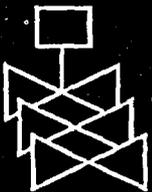
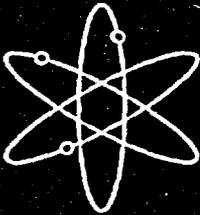
Volume 2

- Digital Instrumentation and Control
- Structural Performance
- The Halden Program
- PRA Methods and Applications

Held at
Bethesda Marriott Hotel
Bethesda, Maryland
October 26-28, 1998

**U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research**

Proceedings prepared by
Brookhaven National Laboratory



AVAILABILITY NOTICE

Availability of Reference Materials Cited in NRC Publications

NRC publications in the NUREG series, NRC regulations, and *Title 10, Energy, of the Code of Federal Regulations*, may be purchased from one of the following sources:

1. The Superintendent of Documents
U.S. Government Printing Office
P.O. Box 37082
Washington, DC 20402-9328
<http://www.access.gpo.gov/su_docs>
202-512-1800
2. The National Technical Information Service
Springfield, VA 22161-0002
<<http://www.ntis.gov/ordernow>>
703-487-4650

The NUREG series comprises (1) brochures (NUREG/BR-XXXX), (2) proceedings of conferences (NUREG/CP-XXXX), (3) reports resulting from international agreements (NUREG/IA-XXXX), (4) technical and administrative reports and books [(NUREG-XXXX) or (NUREG/CR-XXXX)], and (5) compilations of legal decisions and orders of the Commission and Atomic and Safety Licensing Boards and of Office Directors' decisions under Section 2.206 of NRC's regulations (NUREG-XXXX).

A single copy of each NRC draft report is available free, to the extent of supply, upon written request as follows:

Address: Office of the Chief Information Officer
Reproduction and Distribution
Services Section
U.S. Nuclear Regulatory Commission
Washington, DC 20555-0001
E-mail: <DISTRIBUTION@nrc.gov>
Facsimile: 301-415-2289

A portion of NRC regulatory and technical information is available at NRC's World Wide Web site:

<<http://www.nrc.gov>>

All NRC documents released to the public are available for inspection or copying for a fee, in paper, microfiche, or, in some cases, diskette, from the Public Document Room (PDR):

NRC Public Document Room
2120 L Street, N.W., Lower Level
Washington, DC 20555-0001
<<http://www.nrc.gov/NRC/PDR/pdr1.htm>>
1-800-397-4209 or locally 202-634-3273

Microfiche of most NRC documents made publicly available since January 1981 may be found in the Local Public Document Rooms (LPDRs) located in the vicinity of nuclear power plants. The locations of the LPDRs may be obtained from the PDR (see previous paragraph) or through:

<<http://www.nrc.gov/NRC/NUREGS/SR1350V9/1pdr/html>>

Publicly released documents include, to name a few, NUREG-series reports; *Federal Register* notices; applicant, licensee, and vendor documents and correspondence; NRC correspondence and internal memoranda; bulletins and information notices; inspection and investigation reports; licensee event reports; and Commission papers and their attachments.

Documents available from public and special technical libraries include all open literature items, such as books, journal articles, and transactions, *Federal Register* notices, Federal and State legislation, and congressional reports. Such documents as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings may be purchased from their sponsoring organization.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at the NRC Library, Two White Flint North, 11545 Rockville Pike, Rockville, MD 20852-2738. These standards are available in the library for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from—

American National Standards Institute
11 West 42nd Street
New York, NY 10036-8002
<<http://www.ansi.org>>
212-642-4900

DISCLAIMER

Where the papers in these proceedings have been authored by contractors of the United States Government, neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use or the results of such use, of any information,

apparatus, product, or process disclosed in these proceedings, or represents that its use by such third party would not infringe privately owned rights. The views expressed in these proceedings are not necessarily those of the United States Nuclear Regulatory Commission.

Proceedings of the Twenty-Sixth Water Reactor Safety Information Meeting

Volume 2

- Digital Instrumentation and Control
- Structural Performance
- The Halden Program
- PRA Methods and Applications

Held at
Bethesda Marriott Hotel
Bethesda, Maryland
October 26-28, 1998

Manuscript Completed: May 1999
Date Published: June 1999

Compiled by: Susan Monteleone

S. Nesmith, NRC Project Manager

Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555-0001

Proceedings prepared by
Brookhaven National Laboratory



**NUREG/CP-0166, Vol. 2 has been
reproduced from the best available copy.**

ABSTRACT

This three-volume report contains papers presented at the Twenty-Sixth Water Reactor Safety Information Meeting held at the Bethesda Marriott Hotel, Bethesda, Maryland, October 26-28, 1998. The papers are printed in the order of their presentation in each session and describe progress and results of programs in nuclear safety research conducted in this country and abroad. Foreign participation in the meeting included papers presented by researchers from France, Germany, Italy, Japan, Norway, Russia, Sweden and Switzerland. The titles of the papers and the names of the authors have been updated and may differ from those that appeared in the final program of the meeting.

**PROCEEDINGS OF THE
26TH WATER REACTOR SAFETY INFORMATION MEETING**

OCTOBER 26-28, 1998

Published in Three Volumes

GENERAL INDEX

Volume 1

- Plenary Sessions
- Pressure Vessel Research
- Severe Accident Research, Fission Product Behavior
- Nuclear Materials Issues and Health Effects Research
- Materials Integrity Issues

Volume 2

- Digital Instrumentation and Control
- Structural Performance
- The Halden Program
- PRA Methods and Applications

Volume 3

- Thermal Hydraulic Research
- Plant Aging I - Plant Life Management
- High Burn-up Fuel
- Plant Aging II - Cable Aging

REGISTERED ATTENDEES

26TH WATER REACTOR SAFETY MEETING

D. C. AGARWAL
 U.S. DEPT. OF ENERGY
 19901 GERMANTOWN RD. GERMANTOWN
 MD 20585 USA
 Phone: 301 903 3919
 Fax: 301 903 5057
 E-Mail: dulagarwal@hq.doe.gov

R. AMADOR-GARCIA
 COMISION NACIONAL DE SEGURIDAD
 NUCLEAR
 BARRAGAN DR. #779 MEXICO, D.F. 03020
 MEXICO
 Phone: 525-590-8113
 Fax: 525-590-6103
 E-Mail:

F. AMMIRATO
 EPRI NDE CENTER
 1300 HARRIS BLVD. CHARLOTTE NC 28262
 USA
 Phone: 704 547 6129
 Fax: 704 547 6168
 E-Mail: fammirat@epri.com

S. ANGHAE
 U. FLORIDA, DEPT. NUCLEAR &
 RADIOLOGICAL ENG.
 ROOM 202, NUCLEAR SCIENCES CENTER
 GAINESVILLE FL 32611 USA
 Phone: 352 392 1421
 Fax: 352 392 8656
 E-Mail: anghae@ufl.edu

A. R. ANKRUM
 BATTELLE PNNL
 PO BOX 999, K2-28 RICHLAND WA 99352
 USA
 Phone: 509 372 4095
 Fax: 509 372 6242
 E-Mail: ar_ankrum@pnl.gov

W. H. BAMFORD
 WESTINGHOUSE
 PO BOX 355 PITTSBURGH PA 15238 USA
 Phone: 412 374 6515
 Fax: 412 374 6277
 E-Mail: bamfordwh@westinghouse.com

S. BANERJEE
 UNIVERSITY OF CALIFORNIA
 DEPT. OF CHEMICAL ENGINEERING
 SANTA BARBARA CA 93106 USA
 Phone: 805 893 3456
 Fax: 805 893 4731
 E-Mail: banerjee@anemone.ucsb.edu

R. E. BEEDLE
 NUCLEAR ENERGY INSTITUTE
 1776 EYE ST., NW, SUITE 400
 WASHINGTON DC 20006 USA
 Phone: 202 739 8101
 Fax: 202 785 1898
 E-Mail: rb@nei.org

E. BEK
 PJSC MASHINOSTROITELNY ZAVOD
 ELECTROSTAL MOSCOW REGION 144001
 RUSSIA
 Phone: 7 95 7029731
 Fax: 7 95 5750947
 E-Mail:

K. D. BERGERON
 SANDIA NATIONAL LABORATORIES
 PO BOX 5800, DEPT. 8421/MS0739
 ALBUQUERQUE NM 87185-0739 USA
 Phone: 505 844 2507
 Fax: 505 844 8719
 E-Mail: kdberge@sandia.gov

C. E. BEYER
 BATTELLE/PNNL
 BATTELLE BLVD. RICHLAND WA 99352
 USA
 Phone: 509-372-4605
 Fax: 509-372-4439
 E-Mail: carl.beyer@pnl.gov

D. BHARGAVA
 VIRGINIA POWER
 5000 DOMINION BLVD. GLEN ALLEN VA
 23060 USA
 Phone: 804 273 3638
 Fax: 804 293 3448
 E-Mail: divakar_bhargava@vpower.com

M. BILLONE
 ARGONNE NATIONAL LAB
 9700 S. CASS AVE ARGONNE IL
 60439-4839 USA
 Phone: 630 252 7146
 Fax: 630 252 8232
 E-Mail: billone@anl.gov

N. E. BIXLER
 SANDIA NATIONAL LABORATORIES
 P.O. BOX 5800, DEPT. 8421/MS0739
 ALBUQUERQUE NM 87185-0739 USA
 Phone: 505 845 3144
 Fax: 505 844 8719
 E-Mail: nbixler@sandia.gov

J. E. BONDARYK
 ENGINEERING TECHNOLOGY CENTER
 84 SHERMAN ST. CAMBRIDGE MA 02148
 USA
 Phone: 617 864 1944
 Fax: 617 864 1953
 E-Mail: jbondaryk@etc.stinc.com

G. A. BROWN
 AEA TECHNOLOGY
 THOMSON HOUSE, RISLEY WARRINGTON
 CHESHIRE WA3 8AT ENGLAND
 Phone: 44 19 25 254473
 Fax: 44 19 25 254473
 E-Mail: geoff.brown@aeat.co.uk

T. J. BROWN
 SANDIA NATIONAL LABORATORIES
 P.O. BOX 5800 ALBUQUERQUE NM
 87185-0736 USA
 Phone: 505 844 5247
 Fax:
 E-Mail: tbrown@sandia.gov

D. A. BROWNSON
 IDAHO NATIONAL ENGINEERING &
 ENVIRONMENTAL LAB
 PO BOX 1625 IDAHO FALLS ID 83415-3850
 USA
 Phone: 208 526 9460
 Fax: 208 526 2930
 E-Mail: dov@inel.gov

W. T. BRUNSON
 FRAMATOME COGEMA FUELS
 3315 OLD FOREST RD. LYNCHBURG VA
 24503 USA
 Phone: 804 832 2687
 Fax: 804 832 3663
 E-Mail: wbrunson@framatech.com

J. W. BRYANT
 LOCKHEED MARTIN IDAHO
 TECHNOLOGIES CO.
 PO BOX 1625 IDAHO FALLS ID 83415-3114
 USA
 Phone: 208 526 3981
 Fax: 208 526 4902
 E-Mail: bryajw@inel.gov

J. C. BUTLER
 NUCLEAR ENERGY INSTITUTE
 1776 EYE ST., NW, SUITE 400
 WASHINGTON DC 20006 USA
 Phone: 202 739 8000
 Fax: 202 785 1898
 E-Mail: jcb@nei.org

S. T. BYRNE
 ABB
 2000 DAY HILL RD., MC 9483-1903
 WINDSOR CT 06095 USA
 Phone: 860 285 3469
 Fax: 860 285 4232
 E-Mail: stephen.lbyrne@ussev.mail.abb.com

A. L. CAMP
 SANDIA NATIONAL LABORATORIES
 PO BOX 5800, MS 0747 ALBUQUERQUE NM
 87185-0747 USA
 Phone: 505 844 5960
 Fax: 505 844 3321
 E-Mail: alcamp@sandia.gov

J. J. CAREY
 EPRI
 3412 HILLVIEW AVE PALO ALTO CA 94304
 USA
 Phone: 650 855 2105
 Fax: 650 855 7945
 E-Mail: jcarey@epri.com

Y. C. CHI
DEPT. OF NUCLEAR REG., ATOMIC
ENERGY COMM.
67 LANE 144, KEELUNG RD, SEC. 4 TAIPEI
TAIWAN 10690 REP. CHINA
Phone: 886 2 23634180
Fax: 886 2 23635377
E-Mail: chiyc@ccz2.aec.gov.tw

A. B. COHEN
ARGONNE NATIONAL LAB
9700 S. CASS AVE ARGONNE IL
60439-4838 USA
Phone: 630 252 5179
Fax: 630 252 9232
E-Mail: adam.cohen@anl.gov

K. O. COZENS
NUCLEAR ENERGY INSTITUTE
1776 EYE ST., NW, SUITE 400
WASHINGTON DC 20006 USA
Phone: 202 739 8000
Fax: 202 785 1898
E-Mail: koc@nei.org

G. L. DARDEN
VIRGINIA POWER
5000 DOMINION BLVD, IN3S GLEN ALLEN
VA 23060 USA
Phone: 804 273 3497
Fax: 804 273 3543
E-Mail: gary_darden@vapower.com

M. S. DESAI
UNDERWRITERS LAB
12 LABORATORY DRIVE, P.O. BOX 13995
RESEARCH TRIANGLE PARK NC 27709
USA
Phone: 919 549 1610
Fax: 919 547 6110
E-Mail: desaim@ui.com

S. DOROFEEV
RUSSIAN RESEARCH CENTER,
KURCHATOV INSTITUTE
KIRCHATOV SQ. 1 MOSCOW 123182
RUSSIA
Phone:
Fax:
E-Mail:

B. M. DUNN
FRAMATOME TECHNOLOGIES, INC.
OLD FOREST RD. LYNCHBURG VA 24501
USA
Phone: 804 832 2427
Fax:
E-Mail: bdunn@framatech.com

R. C. EVANS
NUCLEAR ENERGY INSTITUTE
1776 EYE ST., NW, SUITE 400
WASHINGTON DC 20006 USA
Phone: 202 739 8000
Fax: 202 785 1898
E-Mail: rce@nei.org

W. G. CHOE
TU ELECTRIC
1601 N. BRYAN ST. DALLAS TX 75201-3411
USA
Phone: 214 812 4371
Fax: 214 812 8687
E-Mail: whee.choe@tuelectric.com

A. S. COHLMeyer
VPA CORPORATION
1768 BUSINESS CENTER DRIVE RESTON
VA 20190 USA
Phone: 703 438 3911
Fax: 703 438 3911
E-Mail:

D. CRAWFORD
ARGONNE NATIONAL LABORATORY
P.O. BOX 2528 IDAHO FALLS ID 83403 USA
Phone: 208 533 7456
Fax: 208 533 7863
E-Mail: doug.crawford@anlw.anl.gov

R. S. DAUM
PENNSYLVANIA STATE UNIVERSITY
231 SACKETT BLDG., DEPT OF NUC ENG
STATE COLLEGE PA 16802 USA
Phone: 814 863 3512
Fax: 814 865 8499
E-Mail: rsd12r@psu.edu

T. L. DICKSON
LOCKHEED MARTIN ENERGY RESEARCH
PO BOX 2008 OAK RIDGE TN 37831 USA
Phone: 423 574 0650
Fax: 423 576 0651
E-Mail: tyd@ornl.gov

R. L. DOTY
PP&L, INC.
2 N. NINTH ST. (GENA93) ALLENTOWN PA
18101 USA
Phone: 610 774 7932
Fax: 610 774 7205
E-Mail: ridoty@papl.com

F. A. DURAN
SANDIA NATIONAL LABORATORIES
PO BOX 5800, MS0747, DEPT 6412
ALBUQUERQUE NM 87185-0747 USA
Phone: 505 844 4495
Fax: 505 844 3321
E-Mail: faduran@sandia.gov

M. L. EYRE
PECO NUCLEAR
965 CHESTERBROOK BLVD., 62A-5 WAYNE
PA 19087-5691 USA
Phone: 610 640 6829
Fax: 610 640 6797
E-Mail: meyre@peco-energy.com

H. M. CHUNG
ARGONNE NATIONAL LAB
9700 S. CASS AVE ARGONNE IL
60439-4838 USA
Phone: 630 252 5111
Fax: 630 252 3604
E-Mail: hee_chung@qmgate.anl.gov

L. CONNOR
DOC-SEARCH ASSOCIATES
PO BOX 34 CABIN JOHN MD 20818 USA
Phone: 301 346 0119
Fax: 503 973 5037
E-Mail: lynnbc@compuserve.com

M. E. CUNNINGHAM
PACIFIC NORTHWEST NATIONAL LAB
P.O. BOX 999 RICHLAND WA 99337 USA
Phone: 509 372 4587
Fax: 509 372 4989
E-Mail: mitch.cunningham@pnl.gov

J. S. DE BOR
DE BOR AND ASSOCIATES, INC.
3630 NO. 21 AVE. ARLINGTON VA 22207
USA
Phone: 703 524 3222
Fax: 703 524 2427
E-Mail: cc001331@mindspring.com

I. DOR
CEA GRENOBLE/DRN/DTP/SMTH
17 RUE DES MARTYRS GRENOBLE CEDEX
9 38054 FRANCE
Phone: 33 4 76885970
Fax: 33 47 6889453
E-Mail: isabelle.dor@cea.fr

J. D. DUNKLEBERGER
NEW YORK STATE HEALTH DEPT.
II UNIVERSITY PLACE ALBANY NY 12203
USA
Phone: 518 458 6458
Fax: 518 458 6434
E-Mail: jd08@health.state.ny.us

F. A. EMERSON
NUCLEAR ENERGY INSTITUTE
1776 EYE ST., NW, SUITE 400
WASHINGTON DC 20006 USA
Phone: 202 739 8000
Fax: 202 785 1898
E-Mail: fae@nei.org

J. A. FORESTER
SANDIA NATIONAL LABORATORIES
PO BOX 5800, MS 0747 ALBUQUERQUE NM
87185-0747 USA
Phone: 505 844 0578
Fax: 505 844 3321
E-Mail: jafores@sandia.gov

I. FRANKL
STOLLER NUCLEAR FUEL/NAC
INTERNATIONAL
485 WASHINGTON AVENUE
PLEASANTVILLE NY 10570 USA
Phone: 914-741-1200
Fax: 914-741-2093
E-Mail: ifrankl@nacintl.com

T. FUKETA
JAPAN ATOMIC ENERGY RESEARCH
INSTITUTE
TOKAI BARAKI 319-1195 JAPAN
Phone: 81 29 282 6386
Fax: 81 29 282 6160
E-Mail: toyo@nerr.tokai.iaeri.go.jp

P. H. GENOA
NUCLEAR ENERGY INSTITUTE
1776 EYE ST., NW, SUITE 400
WASHINGTON DC 20006 USA
Phone: 202 739 6000
Fax: 202 785 1898
E-Mail: phg@nei.org

R. M. GODFREY
AUSTRALIAN NUCLEAR SCIENCE & TECH.
ORG.
EMBASSY OF AUSTRALIA, 1601 MASS.
AVE., NW WASHINGTON DC 20036 USA
Phone: 202 797 3042
Fax: 202 483 5156
E-Mail:

D. F. GRAND
CEA - NUCLEAR REACTORS
DIRECTORATE
17 RUE DES MARTYRS GRENOBLE CEDEX
9 38054 FRANCE
Phone: 33 4 7688 3933
Fax: 33 4 7688 5179
E-Mail: grand@ntp.cea.fr

M. GREGORIC
SLOVENIAN NUCLEAR SAFETY
ADMINISTRATION
VOJKOVA 59 LJUBLJANA SI 01113
SLOVENIA
Phone: 386 61 172 11 00
Fax: 386 61 172 11 99
E-Mail: miroslav.gregoric@rujv.sigov.mail.si

R. O. HARDIES
BGE
1650 CALVERT CLIFFS PKWY LUSBY MD
20732 USA
Phone: 410-495-6577
Fax: 410-492-6577
E-Mail: robert.o.hardies@bge.com

L. HENDRICKS
NUCLEAR ENERGY INSTITUTE
1776 EYE ST., NW, SUITE 400
WASHINGTON DC 20006 USA
Phone: 202 739 8000
Fax: 202 785 1898
E-Mail: bh@nei.org

Y. FUJIKI
TOSHIBA INTERNATIONAL CORP.
175 CURTNER AVENUE SAN JOSE CA USA
Phone: 408-925-6592
Fax: 408-925-4945
E-Mail: yasunobu.fujiki@toshiba.co.jp

F. GANTENBEIN
INSTITUT DE PROTECTION ET DE SURETE
NUCLEAIRE
BP 6 FONTENAY-AUX-ROSES CEDEX
92265 FRANCE
Phone:
Fax:
E-Mail: francase.gantenbein@ipsn.fr

G. GIGGER
WESTINGHOUSE
P.O. BOX 79 WEST MIFFLIN PA 15122 USA
Phone: 412 476 7365
Fax:
E-Mail:

M. GOMOLINSKI
INSTITUT DE PROTECTION ET DE SURETE
NUCLEAIRE
BP 6 FONTENAY-AUX-ROSES 92265
FRANCE
Phone: 146548177
Fax: 146548925
E-Mail: maurice.gomolinski@ipsn.fr

C. GRANDJEAN
INSTITUT DE PROTECTION ET DE SURETE
NUCLEAIRE
CEA CADARACHE ST PAUL LEZ DURANCE
13108 FRANCE
Phone: 33 4 4225 4480
Fax: 33 4 4225 6142
E-Mail: claude.grandjean@ipsn.fr

J. HA
KOREA ATOMIC ENERGY RESEARCH
INSTITUTE
150 DUKJINDONG, YUSUNG-KU TAEJON
305-353 KOREA
Phone: 82 42 8582755
Fax: 82 42 8686374
E-Mail: jha@naram.kaeri.re.kr

J. J. HARTZ
WESTINGHOUSE ELECTRIC
P.O. BOX 355 PITTSBURGH PA 15230 USA
Phone: 412 374 5185
Fax:
E-Mail: hartzjj@westinghouse.com

J.Y. HENRY
CEA/IPSNUDES/SAMS/BASP
BP 6 FONTENAY-AUX-ROSES 92265
FRANCE
Phone: 01 46 54 90 16
Fax: 01 47 46 10 14
E-Mail: jean.yves-henry@ipsn.fr

M. FUJITA
KANSAI ELECTRIC POWER CO., INC.
2001 L ST., NW, SUITE 801 WASHINGTON
DC 20036 USA
Phone: 202 859 1138
Fax: 202 457 0272
E-Mail: mfujita@kansai.com

G. GAUTHIER
CEA/IPSNUDES/SAMS/BASME
BP 6 FONTENAY-AUX-ROSES 92265
FRANCE
Phone: 01 46 54 90 16
Fax: 01 47 46 10 14
E-Mail:

K. T. GILLEN
SANDIA NATIONAL LABORATORY
ORG. 1811 - M/S 1407, P.O. BOX 5800
ALBUQUERQUE NM 87185-1407 USA
Phone: 505 844 7494
Fax: 505 844 9624
E-Mail: ktgille@sandia.gov

A. L. GRAHAM
COUNCIL FOR NUCLEAR SAFETY
PO BOX 7106 CENTURION GAUTENG
00046 SOUTH AFRICA
Phone: 27 12 6635500
Fax: 27 12 6635513
E-Mail: agraham@cns.co.za

M. GREEN
OECD HALDEN REACTOR PROJECT
P.O. BOX 173, N-1751 HALDEN NORWAY
Phone: 47 69212200
Fax: 47 69212201
E-Mail:

B. P. HALLBERT
LOCKHEED-MARTIN
P.O. BOX 1625 IDAHO FALLS ID 83415 USA
Phone: 208 526 9867
Fax:
E-Mail: hallbp@inel.gov

R. C. HARVIL
CONSUMERS ENERGY, PALISADES
NUCLEAR PLANT
27780 BLUE STAR MEMORIAL HWY
COVERT MI 49043 USA
Phone: 616 764 2954
Fax: 616 764 2060
E-Mail:

D. C. HERRELL
MPR ASSOCIATES, INC.
320 KING ST. ALEXANDRIA VA 22314 USA
Phone: 703 519 0200
Fax: 703 519 0220
E-Mail: dherrell@mpr.com

C. HERRERA
CHUBU ELECTRIC POWER CO.
900 17TH ST, NW, STE 1220 WASHINGTON
DC 20008 USA
Phone: 202 775 1960
Fax: 202 331 9256
E-Mail: carolina@chubudc.com

J. C. HIGGINS
BROOKHAVEN NATIONAL LABORATORY
PO BOX 5000, BLDG. 130 UPTON NY
11973-5000 USA
Phone: 516 344 2432
Fax: 516 344 3957
E-Mail: higgins@bnl.gov

J. S. HOLM
SIEMENS POWER CORP.
2101 HORN RAPIDS RD. RICHLAND WA
99352 USA
Phone: 509 375 8142
Fax: 509 375 8775
E-Mail: jerrys_holm@nrfuel.com

T. HSU
VIRGINIA POWER
5000 DOMINION BLVD. GLEN ALLEN VA
23060 USA
Phone:
Fax:
E-Mail:

H. T. HUNTER
LOCKHEED MARTIN ENERGY RESEARCH
PO BOX 2008 OAK RIDGE TN 37831-6362
USA
Phone: 423 576-6297
Fax: 423 574 6182
E-Mail: h30@ornl.gov

J. E. HUTCHINSON
EPRI
1300 HARRIS BLVD. CHARLOTTE NC 28262
USA
Phone: 704 547 6086
Fax: 704 547 6035
E-Mail: jhutchin@epri.com

J.P. C. HUTIN
ELECTRICITE DE FRANCE
DEPT. 1, PLACE PLEYEL ST DENIS CEDEX
93282 FRANCE
Phone: 33 1 43693051
Fax: 33 1 43693495
E-Mail: jean-pierre.hutin@edf.gdf.fr

J. R. IRELAND
LOS ALAMOS NATIONAL LABORATORY
PO BOX 1663, MS F608 LOS ALAMOS NM
87545 USA
Phone: 505 667 4567
Fax: 505 665 5204
E-Mail: john.ireland@lanl.gov

S. K. ISKANDER
OAK RIDGE NATIONAL LABORATORY
MS 6151, BLDG. 45005, P.O. BOX 2008 OAK
RIDGE TN 37831-6151 USA
Phone: 423-574-4468
Fax: 423-574-5118
E-Mail: ski@ornl.gov

R. IWASAKI
NUCLEAR POWER ENGINEERING CORP.
FUJITA KANKO TORANOMON BLDG, 6F
MINATO-KU TOKYO 105-0001 JAPAN
Phone: 81 3 3438 3068
Fax: 81 3 5470 5544
E-Mail:

R. JANATI
DEPT. OF ENVIR. PROT., DIV. OF
NUCLEAR SAFETY
PO BOX 8489, 400 MARKET ST.
HARRISBURG PA 17105 USA
Phone: 717 787 2163
Fax: 717 783 8965
E-Mail: janatrich@91.dep.state.pa.us

J. V. JANERI
UNDERWRITERS LABORATORIES, INC.
12 LABORATORY DR. RESEARCH
TRIANGLE PARK NC 27709 USA
Phone: 919 549 1902
Fax: 919 547 6113
E-Mail: janerj@ul.com

J. JANSKY
BTB-JANSKY GmbH
GERLINGERSTR. 151 LEONBERG 71229
GERMANY
Phone: 07152 41058
Fax: 07152 73868
E-Mail: btbjansky1@aol.com

T-E. JIN
KOREA POWER ENGINEERING CO.
360-9 MABUK-RI, KUSONG-MYON
YONGIN-CITY KYUNG GI-DD 449713
KOREA
Phone: 0331 289 7579
Fax: 0331 289 4517
E-Mail: jinte@ns.kopec.co.kr

B. W. JOHNSON
UNIVERSITY OF VIRGINIA
THORNTON HALL CHARLOTTESVILLE VA
22903-2442 USA
Phone: 804 924 7623
Fax: 804 924 8818
E-Mail: bwj@virginia.edu

W. V. JOHNSTON
RETIRED
2 RUTH LAND DOWNINGTOWN PA 19335
USA
Phone: 610 873 7182
Fax: 610 873 7182
E-Mail: wjohn@nri.com

C. R. JONES
TECHNIDIGM ORG.
13624 HARTSBOURNE DR. GERMANTOWN
MD 20874 USA
Phone: 301-972-2017
Fax: 301-428-9341
E-Mail: tech2000@ix.netcom.com

E. KAPLAR
RUSSIAN RESEARCH CENTER,
KURCHATOV INSTITUTE
KIRCHATOV SQ. 1 MOSCOW 123182
RUSSIA
Phone: 7 095 198 9725
Fax: 7 095 198 1702
E-Mail: asmolov@nsi.kiae.ru

T. M. KARLSEN
OECD HALDEN REACTOR PROJECT
P.O. BOX 173, N-1751 HALDEN NORWAY
Phone: 47 69212200
Fax: 47 69212201
E-Mail:

L. M. KAUFMAN
UNIVERSITY OF VIRGINIA
THORNTON HALL CHARLOTTESVILLE VA
22901 USA
Phone: 804 924 6083
Fax: 804 924 8818
E-Mail: lori@virginia.edu

P. J. KERSTING
KW CONSULTING, INC.
PO BOX 101567 PITTSBURGH PA 15237
USA
Phone: 412 635 7333
Fax: 412 367 2195
E-Mail: paul@kwconsulting.com

H. KIM
COMMONWEALTH EDISON
1400 OPUS DR., STE. 400 DOWNERS
GROVE IL 60515 USA
Phone: 630 663 3072
Fax: 630 663 7181
E-Mail: hak-soo.kim@ucm.com

B. L. KIRK
OAK RIDGE NATIONAL LABORATORY
BLDG. 6025, PO BOX 2008 OAK RIDGE TN
37831-6362 USA
Phone: 423 574 6176
Fax: 423 574 6182
E-Mail: blk@ornl.gov

R. W. KNOLL
FLORIDA POWER CORP.
1022 POWERLINE ROAD CRYSTAL RIVER
FL
Phone:
Fax:
E-Mail:

T. S. KRESS
U.S. NRC/ACRS
102-B NEWRIDGE RD. OAK RIDGE TN
37830 USA
Phone: 423 483 7548
Fax: 423 462 7548
E-Mail: tkress@aol.com

K. F. KUSSMAUL
UNIVERSITY OF STUTTGART
PFAFFENWALDRING 32 STUTTGART
D70569 GERMANY
Phone: 49 711 685 3582
Fax: 49 711 685 2635
E-Mail: kussmaul@mpa.uni-stuttgart.de

C.M. LEE
KOREA POWER ENGINEERING CO.
360-9 MABUK-RI, KUSONG-MYON
YONGIN-CITY KYUNG GI-DO 449713
KOREA
Phone: 0331 289 3579
Fax: 0331 289 4517
E-Mail: cmlee@ns.kopec.co.kr

R. LOFARO
BROOKHAVEN NATIONAL LABORATORY
PO BOX 5000, BLDG. 130 UPTON NY
11973-5000 USA
Phone: 516 344 7191
Fax: 516 344 5569
E-Mail: lofaro@bnl.gov

S. MAJUMDAR
ARGONNE NATIONAL LAB
9700 S. CASS AVE ARGONNE IL
60439-4838 USA
Phone: 630 252 5136
Fax: 630 252 9232
E-Mail: saurin_majumdar@qmgate.anl.gov

P. MARSILI
AGENZIA NAZIONALE PROTEZIONE
AMBIENTS
VIA VITALIANO BRANCANTI 48 ROME
00144 ITALY
Phone:
Fax:
E-Mail:

R. K. McGUIRE
RISK ENGINEERING, INC.
4155 DARLEY AVE, SUITE A BOULDER CO
80303 USA
Phone: 303 499 3000
Fax: 303 499 4850
E-Mail: info@riskeng.com

D. B. MITCHELL
FRAMATOME COGEMA FUELS
3315 OLD FOREST ROAD LYNCHBURG VA
24506-0935 USA
Phone: 804 832 3438
Fax: 804 832 3200
E-Mail: dmitchell@framatech.com

K. KUGIMIYA
MITSUBISHI HEAVY INDUSTRIES
AMERICA, INC.
105 MALL BLVD, EXPO MART 339E
MONROEVILLE PA 15146 USA
Phone: 412 374 7395
Fax: 412 374 7377
E-Mail: keiichi_kugimiya@mhiashq.com

J. A. LAKE
LOCKHEED MARTIN IDAHO
TECHNOLOGIES CO.
P.O. BOX 1625 IDAHO FALLS ID 83415-3860
USA
Phone: 208 526 7670
Fax: 208 526 2930
E-Mail: lakeja@inel.gov

Y. LIU
ARGONNE NATIONAL LABORATORY
9700 S. CASS AVENUE ARGONNE IL 60439
USA
Phone: 630-252-5127
Fax: 630-252-3250
E-Mail: yylin@anl.gov

V. K. LUK
SANDIA NATIONAL LABORATORIES
PO BOX 5800, INS DEPT. 8403
ALBUQUERQUE NM 87185-0744 USA
Phone: 505 844 5498
Fax: 505 844 1648
E-Mail: vlduk@sandia.gov

V. MALOFEEV
RUSSIAN RESEARCH CENTER,
KURCHATOV INSTITUTE
KURCHATOV SQ. 1 MOSCOW 123182
RUSSIA
Phone: 7 095 196 7466
Fax: 7 095 196 1702
E-Mail: malofeev@nsi.kiae.ru

M. MASSOUD
BGE NUCLEAR ENGINEERING UNIT
1650 CALVERT CLIFFS PARKWAY, NEF-1
LUSBY MD 20657 USA
Phone: 410 495 6522
Fax: 410 495 4498
E-Mail: mahmoud.massoud@bge.com

J.C. MELIS
INSTITUT DE PROTECTION ET DE SURETE
NUCLEAIRE
BLDG. 250 CE CADARACHE ST PAUL LEZ
DURANCE 01368 FRANCE
Phone: 33 4 4225 8722
Fax: 33 4 4225 2971
E-Mail: jean-claude.melis@ipsn.fr

D. J. MODEEN
NUCLEAR ENERGY INSTITUTE
1776 EYE ST., NW, SUITE 400
WASHINGTON DC 20006 USA
Phone: 202 739 8000
Fax: 202 785 1898
E-Mail: djm@nei.org

S. KURATA
CHUBU ELECTRIC POWER CO.
900 17TH ST, NW, STE 1220 WASHINGTON
DC 20006 USA
Phone: 202 775 1960
Fax: 202 331 9256
E-Mail: kurata@chubudc.com

C. LECOMTE
INSTITUT DE PROTECTION ET DE SURETE
NUCLEAIRE
BP 6 FONTENAY-AUX-ROSES 92285
FRANCE
Phone: 01 46 54 77 36
Fax: 01 46 54 79 71
E-Mail: catherine.lecomte@ipsn.fr

M. LIVOLANT
INSTITUT DE PROTECTION ET DE SURETE
NUCLEAIRE
BP 6 FONTENAY-AUX-ROSES CEDEX
92265 FRANCE
Phone:
Fax:
E-Mail:

E. S. LYMAN
NUCLEAR CONTROL INSTITUTE
1000 CONNECTICUT AVE., NW, STE 804
WASHINGTON DC 20006 USA
Phone: 202 822 8444
Fax: 202 452 0892
E-Mail: lyman@nci.org

A. MARION
NUCLEAR ENERGY INSTITUTE
1776 EYE ST., NW, SUITE 400
WASHINGTON DC 20006 USA
Phone: 202 739 8000
Fax: 202 785 1898
E-Mail: am@nei.org

B. MAVKO
JOSEF STEFAN INSTITUTE
JAMOVA LJUBLJANA 01000 SLOVENIA
Phone: 386 61 1885330
Fax: 386 61 1612258
E-Mail: borut.mavko@ijs.si

D. W. MILLER
ILLINOIS POWER CO.
P.O. BOX 678 CLINTON IL 61727 USA
Phone: 217-935-8881
Fax: 217-935-4632
E-Mail:

S. MONTELEONE
BROOKHAVEN NATIONAL LABORATORY
BLDG. 130, 32 LEWIS ROAD UPON NY
11973-5000 USA
Phone: 516 344 7235
Fax: 516 344 3957
E-Mail: monteleo@bnl.gov

R. J. MORANTE
BROOKHAVEN NATIONAL LABORATORY
BLDG. 475C UPTON NY 11973-5000 USA
Phone: 516 344 5860
Fax: 516 344 4255
E-Mail: morante@bnl.gov

D. P. MURLAND
SCIENCE & ENGINEERING ASSOCIATES,
INC.
7918 JONES BRANCH DR, SUITE 500
MCLEAN VA 22102 USA
Phone: 703 761 4100
Fax: 703 761 4105
E-Mail:

L. A. NEIMARK
ARGONNE NATIONAL LABORATORY
9700 S. CASS AVE. ARGONNE IL
60439-4838 USA
Phone: 630 252 5177
Fax: 630 252 9232
E-Mail: laneimark@anl.gov

A. NUNEZ-CARRERA
COMISION NACIONAL DE SEGURIDAD
NUCLEAR
BARRAGAN DR. #779 MEXICO, D.F. 03020
MEXICO
Phone: 525-590-5113
Fax: 525-590-6103
E-Mail:

D. J. OSETEK
LOS ALAMOS TECHNICAL ASSOCIATES
BLDG. 1, SUITE 400, 2400 LOUISIANA
BLVD. NE ALBUQUERQUE NM 87110 USA
Phone: 505 880 3407
Fax: 505 880 3560
E-Mail: djosetek@lata.com

K.B. PARK
KOREA ATOMIC ENERGY RESEARCH
INSTITUTW
PO BOX 105, YUSONG DAEJON 305-600
KOREA
Phone: 82 42 8682239
Fax: 82 42 8688990
E-Mail: kbpark2@nanum.kaeri.rc.kr

M. PEZZILLI
ENEA
C.R. CASACCIA VIA ANGUILLA RESE.301
ROME 00060 ITALY
Phone: 39 06 30484197
Fax: 39 06 30486308
E-Mail: pezzilli@casaccia.enea.it

G. A. POTTS
GENERAL ELECTRIC NUCLEAR ENERGY
CASTLE HAYNE RD., MIC K12, PO BOX 780
WILMINGTON NC 28402-0780 USA
Phone: 910 675 5708
Fax: 910 675 6966
E-Mail: gerald.potts@gene.ge.com

J. E. MORONEY
MPR ASSOCIATES, INC.
320 KING ST. ALEXANDRIA VA 22314 USA
Phone: 703 519 0200
Fax: 703 519 0224
E-Mail: jmoroney@mpr.com

R. K. NADER
DUKE ENERGY CORP.
7812 ROCHESTER HWY. SENECA SC
29679 USA
Phone: 864 885 4168
Fax: 864 885 3401
E-Mail: rfnader@duke-energy.com

J. NESTELL
MPR ASSOCIATES, INC.
320 KING STREET ALEXANDRIA VA 22314
USA
Phone: 703 519 0200
Fax: 703 519 0224
E-Mail: jnestell@mpr.com

J. M. O'HARA
BROOKHAVEN NATIONAL LABORATORY
PO BOX 5000, BLDG. 130 UPTON NY
11973-5000 USA
Phone: 516 344 3638
Fax: 516 344 4900
E-Mail: ohara@bnl.gov

F. OWRE
OECD HALDEN REACTOR PROJECT
P.O. BOX 173, N-1751 HALDEN NORWAY
Phone: 47 69212200
Fax: 47 69212201
E-Mail:

W. E. PENNELL
LOCKHEED MARTIN ENERGY RESEARCH
PO BOX 2008 OAK RIDGE TN 37831 USA
Phone: 423 576 8571
Fax: 423 576 0651
E-Mail: pq5@ornl.gov

L. PHILLIPS
UTILITY RESOURCE ASSOCIATES CORP.
1901 RESEARCH BOULEVARD, SUITE 405
ROCKVILLE MD 20850-3164 USA
Phone: 301 294 3069
Fax: 301 294 7879
E-Mail: lepi@urac.com

D. POWERS
NRC/ACRS
7964 SARTAN WAY, NEW ALBUQUERQUE
NM 08709 USA
Phone: 505-821-2735
Fax: 505-821-0245
E-Mail: dapowers.sandia.gov

M. MURATA
NUCLEAR POWER ENGINEERING CORP.
FUJITA KANKO TORANOMON BLDG. 6F
17-1 MINATO-KU TOKYO 105 0001 JAPAN
Phone:
Fax:
E-Mail:

R. K. NANSTAD
OAK RIDGE NATIONAL LABORATORY
PO BOX 2008, MS8151 OAK RIDGE TN
37831-6151 USA
Phone: 423 574 4471
Fax: 423 574 5118
E-Mail: nanstadrk@ornl.gov

H. P. NOURBAKHSH
BROOKHAVEN NATIONAL LABORATORY
PO BOX 5000, BLDG. 130 UPTON NY
11973-5000 USA
Phone: 516-344-5405
Fax: 516 344 3957
E-Mail: nour@bnl.gov

N. ORTIZ
SANDIA NATIONAL LABORATORIES
PO BOX 5800, DEPT. 6400/MS0736
ALBUQUERQUE NM 87185-0736 USA
Phone: 505 844 0577
Fax: 505 844 0955
E-Mail: nortiz@sandia.gov

J. PAPIN
INSTITUT DE PROTECTION ET DE SURETE
NUCLEAIRE
CEA CADARACHE ST PAUL LEZ DURANCE
13108 FRANCE
Phone: 33 4 4225 3463
Fax: 33 4 4225 6143
E-Mail: joelle.papin@ipsn.fr

H. PETTERSSON
VATTENFALL FUEL
FAOK STOCKHOLM S16287 SWEDEN
Phone: 46 87395328
Fax: 46 8128640
E-Mail: hakan@fuel.vattenfall.se

R. POST
NUCLEAR ENERGY INSTITUTE
1776 EYE ST., NW, SUITE 400
WASHINGTON DC 20008 USA
Phone: 202 739 8000
Fax: 202 785 1898
E-Mail: rep@nei.org

J. PUGA
UNESA
FRANCISCO GERYAS 3 MADRID SPAIN
Phone: 34 915674800
Fax: 34 915674988
E-Mail: nuclear@unesa.es

C. PUGH
OAK RIDGE NATIONAL LABORATORY
P.O. BOX 2009, M/S 8063 OAK RIDGE TN
37831 USA
Phone: 423-574-0422
Fax: 423-241-5005
E-Mail: pug@ornl.gov

J. R. RASHID
ANATECH
5435 OBERLIN DRIVE SAN DIEGO CA
92121 USA
Phone: 619-455-6350
Fax: 619-455-1094
E-Mail: joe@anatech.com

N. K. RAY
IDAHO NATIONAL ENG. & ENV. LAB
19901 GERMANTOWN ROAD
GERMANTOWN MD 20874 USA
Phone: 301-903-4126
Fax: 301-903-9902
E-Mail: knr@inel.gov

S. RAY
WESTINGHOUSE ENERGY CENTER
NORTHERN PIKE MONROEVILLE PA 15146
USA
Phone: 412 374 2101
Fax: 412 374 2045
E-Mail: rays@westinghouse.com

P. REGNIER
CEA/PSN/DES/SAMS/BASP
BP 6 FONTENAY-AUX-ROSES 92265
FRANCE
Phone: 01 46 54 90 16
Fax: 01 47 46 10 14
E-Mail:

I. C. RICKARD
ASEA BROWN BOVERI ENGINEERING
SVCS.
200 DAY HILL RD. WINDSOR CT 06095 USA
Phone: 860 285 8678
Fax: 860 285 3253
E-Mail:

J. W. RIVERS
JASON ASSOCIATES CORP.
262 EASTGATE DR., SUITE 335 AIKEN SC
29803 USA
Phone: 803-648-6989
Fax: 803-648-0499
E-Mail: jrivers@scescape.net

G. D. ROBISON
DUKE ENERGY CORP.
526 S. CHURCH ST. CHARLOTTE NC 28202
USA
Phone: 704 382 8685
Fax: 704 382 0368
E-Mail: gdrobiso@duke-energy.com

H. S. ROSENBAUM
EPRI CONSULTANT
917 KENSINGTON DRIVE FREMONT CA
94539 USA
Phone: 510 657 2740
Fax:
E-Mail: hemrosenb@aol.com

T. M. ROSSEEL
OAK RIDGE NATIONAL LABORATORY
PO BOX 2008 OAK RIDGE TN 37631-6158
USA
Phone: 423 574 3380
Fax: 423 574 5118
E-Mail: rosseetm@ornl.gov

J. G. ROYEN
OECD NUCLEAR ENERGY AGENCY
LE SEINE-ST. GERMAIN-12 BLVD. DES ILES
ISSY-LES-MOULINEAUX F92130 FRANCE
Phone: 33 1 4524 1052
Fax: 33 1 4524 1129
E-Mail: jackques.royen@oecd.org

L. P. RUIZ
COMISION NACIONAL DE SEGURIDAD
NUCLEAR
DR. BARRAGAN 779 COL. NARVARTE
MEXICO, D.F. 03020 MEXICO
Phone: 525 590 5054
Fax: 525 590 7508
E-Mail: gsn1@servidor.uncm.mx

A. RYDL
NUCLEAR RESEARCH INSTITUTE REZ
25068 REZ NEAR PRAGUE REZ 25068
CZECH REPUBLIC
Phone: 420 2666172471
Fax: 420 220941029
E-Mail: ryd@nri.cz

O. SANDERVAG
SWEDISH NUCLEAR POWER
INSPECTORATE
STOCKHOLM 10658 SWEDEN
Phone: 46 8 6988463
Fax: 46 8 6619086
E-Mail: oddbjorn@ski.se

P. A. SCHEINERT
BETTIS ATOMIC POWER LABORATORY
PO BOX 79 WEST MIFFLIN PA 15521-0079
USA
Phone: 412 476 5974
Fax: 412 476 6937
E-Mail:

C. S. SCHLASEMAN
MPR ASSOCIATES, INC.
320 KING STREET ALEXANDRIA VA 22314
USA
Phone: 703 519 0200
Fax: 703 519 0224
E-Mail: cschlaseman@mpr.com

F. K. SCHMITZ
INSTITUT DE PROTECTION ET DE SURETE
NUCLEAIRE
CEA CADARACHE ST PAUL LEZ DURANCE
13108 FRANCE
Phone: 33 4 4225 7035
Fax: 33 4 4225 2971
E-Mail: franz.schmitz@ipsn.fr

M. SCHWARZ
INSTITUT DE PROTECTION ET DE SURETE
NUCLEAIRE
CENTRE D'ETUDES DE CADARACHE, BAT.
250 ST PAUL LEZ DURANCE 13108
FRANCE
Phone: 33 4 4225 7748
Fax: 33 4 4225 2971

E. SCOTT DE MARTINVILLE
C E A
80, GAL LECLERC FONTENAY AUX ROSES
92265 FRANCE
Phone: 33 1 46548202
Fax: 33 1 46543264
E-Mail:

S. Y. SHIM
ATOMIC ENERGY CONTROL BOARD
280 SLATER ST. OTTAWA ONTARIO
K1P5S9 CANADA
Phone: 613 947 1443
Fax: 613 995 2125
E-Mail: shim.s@atomcon.gc.ca

F. A. SIMONEN
PACIFIC NORTHWEST NATIONAL
LABORATORY
P.O. BOX 999 RICHLAND WA 99352 USA
Phone: 509-375-2087
Fax: 509-375-3614
E-Mail: fa_simonen@pnl.gov

B.P. SINGH
JUPITOR CORPORATION
2730 UNIVERSITY BLVD. W, STE 900
WHEATON MD 20902 USA
Phone: 301 946 8088
Fax: 301 946 6539
E-Mail: bhupinder.singh@hq.doe.gov

T. SIVERTSEN
OECD HALDEN REACTOR PROJECT
P.O. BOX 173, N-1751 HALDEN NORWAY
Phone: 47 69212200
Fax: 47 69212201
E-Mail:

W. H. SLAGLE
WESTINGHOUSE ELECTRIC
P.O. BOX 355 PITTSBURGH PA 15230 USA
Phone: 412 374 2088
Fax: 412 374 2045
E-Mail: slaglewh@westinghouse.com

L. SLEGERS
SIEMENS
POSTFACH 101063 OFFENBACH D63010
GERMANY
Phone:
Fax:
E-Mail:

A. SMIRNOV
RIAR
ULJANOVSK, DIMITROVGRAD RUSSIA
Phone: 7 84235 32350
Fax: 7 84235 64163
E-Mail:

V. SMIRNOV
RIAR
ULJANOVSK, DIMITROVGRAD RUSSIA
Phone: 7 84235 32350
Fax: 7 84235 64163
E-Mail:

C. L. SMITH
INEEL
2525 FREEMONT IDAHO FALLS ID 83415
USA
Phone: 208 526 9804
Fax:
E-Mail: cts2@inel.gov

G. P. SMITH
ABB COMBUSTION ENGINEERING
NUCLEAR POWER
2000 DAY HILL ROAD WINDSOR CT
06095-0500 USA
Phone: 860-687-8070
Fax: 860-687-8051
E-Mail:

P. SOO
BROOKHAVEN NATIONAL LABORATORY
PO BOX 5000, BLDG. 130 UPTON NY
11973-5000 USA
Phone: 516 344 4094
Fax: 516 344 5569
E-Mail: soo@bnl.gov

S. SPALJ
FER-ZAGREB
PRISAVLJE 8 ZAGREB CROATIA
Phone: 385-16129994
Fax: 385-16129890
E-Mail: srojans.spalj@fer.hr

K. SPANG
INGEMANSSON TECHNOLOGY AB
SWEDEN
Phone: 46 31 774 7401
Fax: 46 31 774 7474
E-Mail: kjell.spang@ingemansson.se

N. N. SRINIVAS
DETROIT EDISON
2000 SECOND AVE, WSC H-60 DETROIT MI
48226 USA
Phone: 313 897 1198
Fax: 313 897 1440
E-Mail: srinivasn@dte.com

R. G. STARCK
MPR ASSOCIATES, INC.
320 KING ST. ALEXANDRIA VA 22314 USA
Phone: 703 519 0200
Fax: 703 519 0224
E-Mail:

J. STONE
MPR ASSOCIATES, INC.
320 KING ST. ALEXANDRIA VA 22314 USA
Phone:
Fax:
E-Mail:

P. STOREY
HSE
ST. PETERS HOUSE BOOTLE LIVERPOOL
L203PT UK
Phone: 44 1519514172
Fax: 44 1519513942
E-Mail: peter.storey@HSE.gov.uk

Y. TAKAHASHI
TOKYO ELECTRIC POWER CO.
1-3-1 UCHISAWAI CHO CHIYODAKU
TOKYO 100-0011 JAPAN
Phone: 81 34216 4951
Fax: 81 33596 8571
E-Mail: to560565@pmail.tepco.co.jp

T. TAMINAMI
TOKYO ELECTRIC POWER CO.
1901 L ST, NW, STE 720 WASHINGTON DC
20036 USA
Phone: 202 457 0790
Fax: 202 457 0810
E-Mail: taminami@tepco.com

J. H. TAYLOR
BROOKHAVEN NATIONAL LABORATORY
PO BOX 5000, BLDG. 130 UPTON NY
11973-5000 USA
Phone: 516 344 7005
Fax: 516 344 3957
E-Mail: jtaylor@bnl.gov

V. H. TESCHENDORFF
GESELLSCHAFT FUR ANLAGEN UND
REAKTORSICHERHEIT
FORSCHUNGSGELANDA GARCHING
D85748 GERMANY
Phone: 49 89 32004423
Fax: 49 89 32004599
E-Mail: tes@gra.de

H. O. TEZEL
ATOMIC ENERGY CONTROL BOARD
280 SLATER STREET ONTARIO K1P5S9
CANADA
Phone: 613 995 3896
Fax:
E-Mail: tezel.h@atomcon.gc.ca

H. D. THORNBURG
CONSULTANT
901 S. WARFIELD DR. MT. AIRY MD 21771
USA
Phone: 301 831 7328
Fax: 301 829 0874
E-Mail: matt@erols.com

G. J. TOMAN
NUTHERM INTERNATIONAL, INC.
501 SO. 11 ST MT VERNON IL 62864 USA
Phone: 618 244 6000
Fax: 618 244 6641
E-Mail: nutherm@midwest.net

R. L. TREGONIN
NAVAL SURFACE WARFARE CENTER
9500 MACARTHUR BLVD. WEST
BETHESDA MD 20817 USA
Phone: 301-227-5145
Fax: 301-227-5548
E-Mail: tregonin@metels.dl.navy.mil

S. TSURUMAKI
NUCLEAR POWER ENGINEERING CORP.
SHUWA-KAMIYACHO BLDG., 2F 3-13, 4
CHOME MINATO-KU TOKYO JAPAN
Phone: 81 3 3434 4551
Fax: 81 3 3434 9487
E-Mail:

A. C. UPTON
UMDNJ-RWJ MEDICAL SCHOOL
170 FRELINGHUYSEN RD. PISCATAWAY
NJ 08854 USA
Phone: 732 445 0795
Fax: 732 445 0959
E-Mail: acupton@ehsi.rutgers.edu

R. A. VALENTIN
ARGONNE NATIONAL LABORATORY
9700 S. CASS AVE., BLDG. 308 ARGONNE
IL 60439 USA
Phone: 630 252 4483
Fax: 630 252 3250
E-Mail: richv@anl.gov

K. K. VALTONEN
RADIATION & NUCLEAR SAFETY
AUTHORITY
PO BOX 14 HELSINKI 00881 FINLAND
Phone: 358 9 759 88 331
Fax: 358 9 759 88 382
E-Mail: keijo.valtonen@stuk.fi

J. L. VILLADONIGA
CONSEJO DE SEGURIDAD NUCLEAR
JUSTO DORADO, 11 MADRID 28040 SPAIN
Phone: 34 91 3460240
Fax: 34 91 3460588
E-Mail: jlv@csn.es

C. VITANZA
OECD HALDEN REACTOR PROJECT
OS ALLE 13, PO BOX 173 HALDEN 01751
NORWAY
Phone: 47 69212200
Fax: 47 69212201
E-Mail: carlo.vitanza@hrp.no

L. WARNKEN
SIEMENS KWU NLE
PO BOX 2032 ERLANGEN BAYERN 91050
GERMANY
Phone: 49 91 3118 3336
Fax: 49 91 3118 6362
E-Mail: lueder.warnken@er11.siemens.de

L. E. WILLERTZ
PP&L, INC.
2 NO. NINTH ST., GENA62 ALLENTOWN PA
18101 USA
Phone: 610 774 7646
Fax: 610 774 7830
E-Mail: lewillertz@papa.com

R. YANG
EPRI
3412 HILLVIEW AVE. PALO ALTO CA 94024
USA
Phone: 650 855 2481
Fax: 650 855 1026
E-Mail: ryang@epri.com

K. K. YOON
FRAMATOME TECHNOLOGIES
3315 OLD FOREST RD. LYNCHBURG VA
24506-0935 USA
Phone: 804 832 3280
Fax:
E-Mail:

M. VILLARAN
BROOKHAVEN NATIONAL LABORATORY
PO BOX 5000, BLDG. 130 UPTON NY
11973-5000 USA
Phone: 516 344 3833
Fax: 516 344 5569
E-Mail: villaran@bnl.gov

R. VON ROHR
INST. OF PROCESS ENGINEERING, ETH
ZURICH
SONNEGGSTRASSE 3, PO BOX ZURICH
CH 8092 SWITZERLAND
Phone: 4116322488
Fax: 4116321141
E-Mail: vorrohr@ivuk.mavt.ethz.ch

R. A. WEINER
KW CONSULTING, INC.
PO BOX 101567 PITTSBURGH PA 15237
USA
Phone: 412 635 7732
Fax: 412 687 3965
E-Mail: bob@kwconsulting.com

D. H. WILLIAMSON
SAIC
Phone: 703-827-4896
Fax:
E-Mail:

P.C. YEY
DEPT. OF NUCLEAR REG., ATOMIC
ENERGY COMM.
67 LANE 144, KEELUNG RD, SEC. 4 TAIPEI
TAIWAN 10660 REP. CHINA
Phone: 886 2 23634180
Fax: 886 2 23635377
E-Mail: pcyeh@aec.gov.tw

D. ZANOBETTI
UNIV. OF BOLOGNA
VIALE RISORGIMENTO 2 BOLOGNA I40136
ITALY
Phone: 39 051 6443471
Fax: 39 051 6443470
E-Mail: dino.zanobetti@mail.ing.unibo.it

G. L. VINE
EPRI
2000 L. ST. NW, SUITE 805 WASHINGTON
DC 20036 USA
Phone: 202-293-6347
Fax: 202-293-2697
E-Mail: gvine@epri.com

N. WAECKEL
ELECTRICITE DE FRANCE SEPTEN
12-14 AV DUTRIEVOS VILLEURBANNE
69628 FRANCE
Phone: 33 4 7282 7571
Fax: 33 4 7282 7713
E-Mail: nicolas.waeckel@edf.gdf.fr

W. WIESENACK
OECD HALDEN REACTOR PROJECT
P.O. BOX 173, N-1751 HALDEN NORWAY
Phone: 47 69212200
Fax: 47 69212201
E-Mail:

R. T. WOOD
OAK RIDGE NATIONAL LABORATORY
PO BOX 2008, BLDG. 3500, MS6010 OAK
RIDGE TN 37831-6010 USA
Phone: 423 574 5578
Fax: 423 576 8380
E-Mail: woodrt@ornl.gov

T. YONOMOTO
JAERI, DEPT. OF REACTOR SAFETY
ENGR.
SHIRAKATA TOKAI IBARAKI 319-11 JAPAN
Phone: 81 29 2825262
Fax: 81 29 2826728
E-Mail: yonomoto@lstf3.tokai.jaeri.go.jp

G. L. ZIGLER
ITS CORPORATION
6000 UPTOWN BLVD., NE, STE 300
ALBUQUERQUE NM 87123 USA
Phone: 505 872 1084
Fax: 505 872 0233
E-Mail: gzigler@itsc.com

**Proceedings of the
26th Water Reactor Safety Information Meeting
October 26-28, 1998**

Contents - Volume 2

| | <u>Page</u> |
|----------------------------|-------------|
| Abstract | iii |
| General Index | v |
| Registered Attendees | vii |

**Digital Instrumentation and Control
J. Calvert, Chair**

| | |
|--|-----------|
| Issues in the Design of Human-Systems Interfaces to Digital Systems | 1 |
| <p style="margin-left: 40px;">J. O'Hara, W. Stubler (BNL), J. Kramer (NRC)</p> | |
| The Importance of Fault Detection Coverage in Safety Critical Systems | 5 |
| <p style="margin-left: 40px;">L. Kaufman, B. Johnson (U. Virginia)</p> | |
| Comparing the Safety Criteria of IEC 61508 and UL 1998 | 29 |
| <p style="margin-left: 40px;">J. Janeri, H. Cox, B. Godwin (Underwriters Laboratories, Inc.)</p> | |
| Putting Principles into Practise: The Formal Development of a Theorem Prover | 39 |
| <p style="margin-left: 40px;">T. Sivertsen (OECD Halden Reactor Project)</p> | |

**Structural Performance
A. Murphy, Chair**

| | |
|---|------------|
| Re-evaluation of Regulatory Guidance on Model Response Combination Methods for Seismic Response Spectrum Analysis | 61 |
| <p style="margin-left: 40px;">R. Morante, Y. Wang (BNL), W. Norris (NRC)</p> | |
| Post Test Analysis of a PCCV Model Subjected to Beyond-Design-Basis Earthquake Simulations | 81 |
| <p style="margin-left: 40px;">R. James, Y. Rashid (Anatech), J. Cherry (SNL), N. Chokshi (NRC)</p> | |
| Steel Containment Vessel Model Test: Results and Post-test Analysis | 107 |
| <p style="margin-left: 40px;">V. Luk, J. Ludwigsen, M. Hessheimer (SNL), K. Komine, M. Iriyama (NUPEC), T. Matsumoto (Hitachi), J. Costello (NRC)</p> | |

| | <u>Page</u> |
|---|-------------|
| New Seismic Design Spectra for Nuclear Power Plants | 127 |
| R. McGuire (Risk Engineering), W. Silva (Pacific Engineering & Analysis), R. Kenneally (NRC) | |
| Damage Mechanics-Based Assessment of Time-Dependent Structural Deterioration and Reliability | 141 |
| B. Ellingwood (Johns Hopkins U.), B. Bhattacharya (ABS) | |
| Feasibility Study on the Use of Ultrasonic Technology on Embedded Corrosion Detection of Nuclear Containment Units, Phase II | 157 |
| J. Rudzinsky, M. Conti, J. Bondaryk (Cambridge Acoustical Associates) | |
| The Halden Program J. Persensky, Chair | |
| Overview of the OECD Halden Reactor Project | 183 |
| C. Vitanza (OECD Halden) | |
| The OECD Halden Reactor Project Fuels Testing Programme: Methods, Selected Results and Plans | 195 |
| W. Wiesenack, T. Tverberg (OECD Halden Reactor Project) | |
| Achievements and Further Plans for the OECD Halden Reactor Project Materials Programme | 205 |
| T. Karlsen (OECD Halden Reactor Project) | |
| Achievements and Further Plans for the OECD Halden Reactor Project Man-Machine Systems Program | 213 |
| F. Owre (OECD Halden Reactor Project) | |
| Overview and Results from the Human Error Analysis Project 1997-1998 | 231 |
| P. Braarud, et al. (OECD Halden Reactor Project) | |
| Prospects on Combining Software Quality Assurance Techniques | 245 |
| T. Sivertsen (OECD Halden Reactor Project) | |
| Human Factors Engineering and Control Room Design Using a Virtual Reality-based Tool for Design and Testing | 267 |
| M. Louka, C. Holmstrom, F. Owre (OECD Halden Reactor Project) | |

Full Range Signal Validation of PWR Plant Data and Fast Transient Classification Applied in Alarm Handling Using Neuro-Fuzzy Models 283
P. Fantoni, D. Roverso, F. Owre (OECD Halden Reactor Project)

PRA Methods and Applications
N. Siu, Chair

Discussion of Comments from a Peer Review of a Technique for Human Event Analysis (ATHEANA) 303
J. Forester (SNL), A. Ramey-Smith (NRC), D. Bley (Buttonwood Consulting),
A. Kolaczowski, S. Cooper (SAIC), J. Wreathall (John Wreathall & Co.)

Incorporating Aging Effects into Probabilistic Risk Assessment 317
C. Smith, V. Shah (INEEL), G. Apostolakis, T.-M. Kao (MIT)

Accident Sequence Precursor Program Large Early Release Frequency Model Development 343
D. Brownson (INEEL), T. Brown, F. Duran, J. Gregory (SNL),
E. Rodrick (NRC)

Perspectives on Needed Scope and Level of Detail for a PRA Standard 355
M. Drouin, et al. (NRC), D. Bley (Buttonwood Consulting), R. Budnitz (FRA),
A. Camp, J. Forester, D. Whitehead (SNL), A. Kolaczowski, J. Lachance (SAIC),
J. Lehner, W. Pratt (BNL)

NRC Support for the Kalinin (VVER) Probabilistic Risk Assessment 367
D. Bley (Buttonwood Consulting), D. Diamond, et al. (BNL),
D. Johnson (PLG, Inc.), A. Szukiewicz, et al. (NRC)

Risk Comparison of Performing Scheduled Maintenance at Power Vs. During Shutdown 375
A. Buslik (NRC), P. Samanta (BNL), B. Staple (SNL)

ISSUES IN THE DESIGN OF HUMAN-SYSTEMS INTERFACES TO DIGITAL SYSTEMS

**John O'Hara and William Stubler
Brookhaven National Laboratory
Upton, New York**

**Joel Kramer
U.S. Nuclear Regulatory Commission
Washington, DC**

The U.S. Nuclear Regulatory Commission's (NRC) human factors engineering (HFE) design review guidance is described in:

- NUREG-0800, Chapter 18 of the Standard Review Plan (NRC, 1996),
- NUREG-0711, Human Factors Engineering Program Review Model (O'Hara, Higgins, Stubler, Goodman, Eckenrode, Bongarra, and Galletti, 1994), and
- NUREG-0700, Revision 1, Human-System Interface Design Review Guideline (O'Hara, Brown, Stubler, Wachtel, and Persensky, 1996).

While the NUREGs -0800 and -0711 mainly address the process aspects of HFE considerations, NUREG-0700 addresses the detailed implementation of a human-system interface (HSI) design.

In the development of NUREG-0700, Rev 1, several topics were identified as "gaps" because there was an insufficient technical basis upon which to develop guidance (O'Hara, 1994). One gap was hybrid HSIs; i.e., HSIs that result from the combination of digital and traditional HSI technologies. New demands may be imposed on personnel for the operation and maintenance of these systems. These demands may result from many factors including: characteristics of the new technologies, characteristics of the mixture of new and traditional technologies, the process by which the hybrid HSI is developed and implemented, and the way in which personnel are prepared to use the hybrid HSI.

The NRC is currently sponsoring research to (1) better define the effects of hybrid HSIs on personnel performance and plant safety; and (2) develop HFE guidance to support safety reviews in the event that a review of plant modifications involving a safety-significant aspect of HSIs is necessary. This guidance will be integrated into existing regulatory guidance documents and will be used to provide the NRC staff with the technical basis to help ensure that the modifications or HSI designs do not compromise safety.

HSI technology changes and their potential effects on personnel performance were identified based upon published literature, interviews with designers and subject matter experts, and plant visits (O'Hara, Stubler, and Higgins, 1996). The topics were evaluated with respect to their potential safety significance (Stubler, Higgins, and O'Hara, 1996). One topic found to be potentially significant to safety and selected for the development of HFE guidance was Design Analysis, Evaluation, and Implementation of Hybrid HSIs.

This topic addresses analyses and evaluations conducted during the design of upgrades and the way upgrades are introduced into the HSI and incorporated into plant operating practices. Important considerations included the effects upon personnel of temporary and changing HSI configurations, which result from the installation of HSI upgrades. Additional considerations include training and personnel acceptance of HSI changes. Thus, the topic addresses the life cycle of an HSI upgrade from initial planning through design, evaluation, and installation.

With regard to its application to hybrid HSIs in the context of plant modifications, the existing guidance is also limited in an additional way. While the guidance provides for tailoring of the review methods and criteria to the unique circumstances of an individual review, no guidance is available to assist in the identification of the process elements and criteria that are necessary. The extent of plant modifications can range significantly, e.g., for a replacement "in-kind" of a single HSI component to an extensive control room modification from analog to digital technology. Thus, when and where to apply that guidance that is available needs to be addressed.

The objective of the phase of the research that is reported in this paper was to develop human factors review guidance addressing the process by which hybrid HSIs are developed, implemented, and integrated into plant operations. To support this objective, several tasks were performed including:

- Development of a technical basis using human performance research and analyses that are relevant to upgrades,
- Development of HFE review guidelines in a format that is consistent with existing NRC review guidance, and
- Identification of remaining issues for which research results were insufficient to support the development of NRC review guidance.

The status of each will be briefly addressed below (see Stubler, O'Hara, and Higgins, in preparation, for additional detail).

Technical information related to system development and modification was reviewed in order to identify the effects of upgrades on personnel performance. The technical information included basic HFE literature, HFE literature pertaining to complex human-machine systems, and industry experience gained from site visits, interviews, and literature. In addition to performance effects, the types of knowledge and skills that are needed to adapted to an upgrade were identified.

This information was used to develop a characterization framework for describing key characteristics of hybrid HSIs that are important to HFE reviews. This information also served as the technical basis upon which design review guidelines were developed. The NUREG-0700 guidance development methodology was used to convert this technical basis into technically valid review guidance (O'Hara, Brown, and Nasta, 1996).

The guidance addressed the design process and was organized according to the 10 review element of the NUREG-0711:

- HFE Program Management (Element 1)
- Operating Experience Review (OER) (Element 2)
- Functional Requirements Analysis And Allocation (Element 3)
- Task Analysis (Element 4)
- Staffing (Element 5)
- Human Reliability Analysis (Element 6)
- Human-System Interface Design (Element 7)
- Procedure Development (Element 8)
- Training Program Development (Element 9)
- Human Factors Verification And Validation (Element 10).

Within each element, the guidance was further organized into four categories. The first category described the conditions under which the particular NUREG-0711 element is relevant to the review of upgrades. The second category included guidance from the NUREG-0711 that was modified to focus on characteristics and considerations that are relevant upgrades. The third category included guidance that was specifically relevant to upgrades but did not currently appear in the NUREG-0711. The fourth category included considerations that have potential applications beyond upgrades and are possible additions to the more general guidance of the NUREG-0711.

In the course of the guidance development process, several human performance issues associated with upgrades were identified that could not be addressed with the available technical basis. They represent topics for which further research is necessary. These issues include:

- The role of HSI consistency as applied to traditional and digital HSIs
- The effects of HSI design on crew coordination and cooperation
- The role of training in HSI skills
- The effects of the installation process for HSI upgrades upon personnel performance
- Personnel acceptance of upgrades.

In conclusion, design review guidance addressing the design, evaluation, and implementation considerations of HSI upgrades has been developed. This guidance complements the design review guidance that was already developed in other phases of the project to address the characteristics associated with specific technologies such as soft controls (Stubler and O'Hara, in preparation), advanced information systems (O'Hara and Higgins, in preparation), computer-based procedures (O'Hara, Higgins, and Stubler, in preparation), and digital system maintenance (Stubler and Higgins, in preparation).

All of the guidance was peer reviewed and revised accordingly. The guidance documents are expected to be published in 1999.

References

NRC (1996). Standard Review Plan (NUREG-0800, Draft for Comment). Washington, D.C.: U.S. Nuclear Regulatory Commission.

O'Hara, J. (1994). *Advanced Human System Interface Design Review Guideline* (NUREG/CR-5908). Washington, D.C.: U.S. Nuclear Regulatory Commission.

O'Hara, J., Brown, W., and Nasta, K. (1996). *Development of NUREG, 0700, Revision 1* (BNL Technical Report L-1317-2-12/96). Upton, New York: Brookhaven National Laboratory.

O'Hara, J., Brown, W., Stubler, W., Wachtel, J., and Persensky, J. (1996). *Human-system interface design review guideline* (NUREG-0700, Rev. 1). Washington, D.C.: U.S. Nuclear Regulatory Commission.

O'Hara, J. and Higgins, J. (in preparation). *Advanced Information Systems: Review Guidance and Technical Basis* (draft report). Washington, D.C.: U.S. Nuclear Regulatory Commission.

O'Hara, J., Higgins, J., and Stubler, W. (in preparation). *Computer-Based Procedure Systems: Review Guidance and Technical Basis* (draft report) Washington, D.C.: U.S. Nuclear Regulatory Commission.

O'Hara, J., Higgins, J., Stubler, W., Goodman, C., Eckenrode, R., Bongarra, J., and Galletti, G. (1994). *Human Factors Engineering Program Review Model* (NUREG-0711). U.S. Nuclear Regulatory Commission, Washington, D.C.

Stubler, W., Higgins, J., and O'Hara, J. (1996). *Evaluation of the Potential Safety Significance of Hybrid Human-system Interface Topics* (BNL Technical Report J6012-T2-6/96). Upton, New York: Brookhaven National Laboratory.

O'Hara, J., Stubler, W., and Higgins, J. (1996). *Hybrid Human System Interfaces: Human Factors Considerations* (BNL Report J6012-T1-4/96). Upton, New York: Brookhaven National Laboratory.

O'Hara, J., Stubler, W. and Kramer, J. (1997). Addressing the Human Factors Issues Associated With in Control Room Modifications. In *Proceedings of the 25th Water Reactor Safety Information Meeting* (NUREG-CP 0162). Washington, D.C.: U.S. Nuclear Regulatory Commission.

Stubler, W. and Higgins, J. (in preparation). *Maintenance of Digital Systems: Review Guidance and Technical Basis* (draft report). Upton, New York: Brookhaven National Laboratory.

Stubler, W. and O'Hara, J. (in preparation). *Soft Controls: Review Guidance and Technical Basis* (draft report). Washington, D.C.: U.S. Nuclear Regulatory Commission.

Stubler, W., O'Hara, J., and Higgins, J. (in preparation). *Human-System Interface and Plant Modernization Process: Review Guidance and Technical Basis* (draft report). Washington, D.C.: U.S. Nuclear Regulatory Commission.

The Importance of Fault Detection Coverage in Safety Critical Systems

Lori M. Kaufman • University of Virginia • Charlottesville

Barry W. Johnson • University of Virginia • Charlottesville

Index Terms: Fault coverage, Coverage estimation, Fault modeling

Abstract: Fault coverage is an important parameter in measuring system dependability. It can be derived via analytic models or statistical estimation. The analytical models represent the fault behavior and are embedded within the overall system model, and generally solved using behavioral description. Statistical estimators use data collected from physical models to derive coverage estimates. The estimates are derived using variance reduction techniques that require a priori knowledge of the distribution of the system's fault data. If the assumed distribution differs from the actual distribution for the fault data, then the accuracy of the coverage estimate is questionable.

1. Introduction

The sensitivity of dependability metrics to slight variations in fault coverage is well documented [6], [16], [19]. Specifically, a small change in coverage can result in great variations in these metrics. Therefore, it is imperative that an accurate estimate of fault coverage be made. It is the purpose of this paper to survey the various methods that are currently used to model and to estimate fault coverage.

There are both mathematical and qualitative expressions for fault coverage. The mathematical definition is that fault coverage, C , is the conditional probability that a system recovers given that a fault has occurred [6]. It is written as

$$C = P(\text{fault processed correctly} \mid \text{fault existence}) \quad (1)$$

Qualitatively, coverage is a measure of the system's ability to detect, locate, contain and recover from the presence of a fault. There are four primary types of fault coverage available: (1) fault detection coverage; (2) fault location coverage; (3) fault containment coverage; and (4) fault recovery coverage. Thus, the term *fault processed correctly* implies at least one of the four coverage types. A more detailed description of the fault coverage types follows.

Fault detection coverage is the system's ability to detect a fault. For example, a typical system requirement is that a certain percentage of all faults must be detected. The fault detection coverage is then a measure of the system's ability to meet the desired fault detection requirement. Fault location coverage measures a system's ability to locate the cause of faults. A typical system requirement is that faults within replaceable components must be located. Hence, fault location coverage is a measure of the success with which such faults are located. Fault containment coverage measures a system's ability to contain faults within a predefined boundary. For example, if a fault in a sub-system is detected and located, then preventing the effects of the fault from propagating in the system is a measure of fault containment coverage. Finally, fault recovery coverage measures the system's ability to automatically recover from faults and to maintain the correct operation. If a system is required to possess a high fault recovery coverage, then it must also possess high fault detection, fault location and fault containment coverages [19].

The type of coverage required is highly application specific. For example, fail-safe systems require specific knowledge of the fault detection coverage. Conversely, highly-reliable systems that use sparing techniques [19] require knowledge of the fault recovery coverage. Regardless of the type of coverage information that is required by a system, the methodology used to estimate the coverage parameter is the

same. Throughout the remainder of this paper, fault coverage is defined to mean any of the four fault coverage categories that are required for a given application.

Fault coverage is examined in two distinct ways: (1) coverage modeling and (2) parameter estimation. As its name implies, fault coverage modeling is a development of a model for the response of a component to the occurrence of a fault. Parameter estimation is needed for values that are required by coverage models. The parameters can be estimated by inserting faults into a given system prototype/model and collecting the required data. There are three primary types of models used in examining coverage [7]:

- (1) *axiomatic models*: analytical models used to model structure and the dependability and/or performance behavior of a system [3].
- (2) *empirical models*: statistical models used to model complex and detailed descriptions of a system's parameter(s) using data collected from physical models.
- (3) *physical models*: prototypes that actually implement the hardware and/or the software of an actual system.

These models allow for different levels of abstraction to be identified during testing as shown in Figure 1. The axiomatic models measure the dependability metrics. In these models, the behavior of a

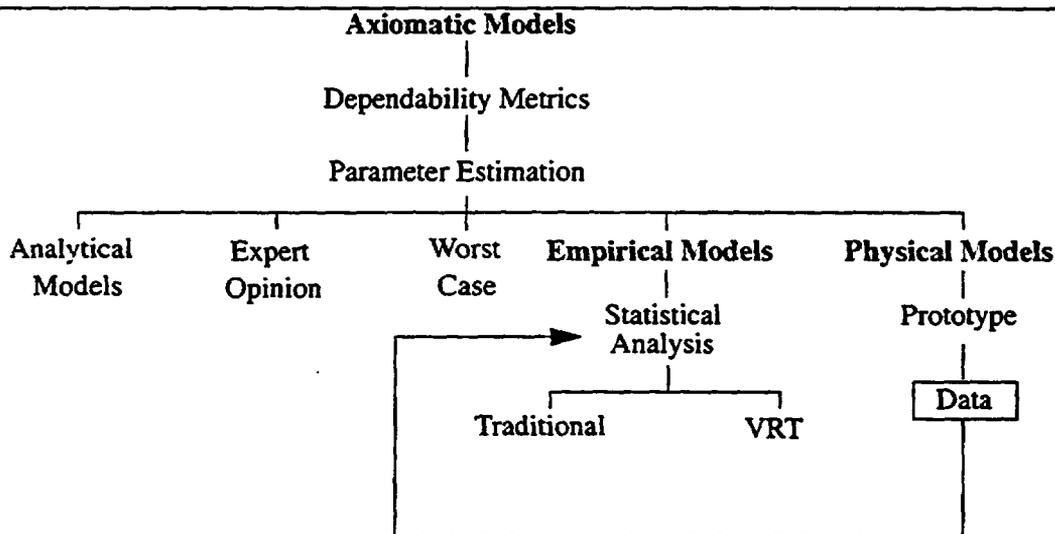


Figure 1. Coverage modeling hierarchy

faulty component is represented and fault coverage is a parameter. The parameter values used are approximations derived from expert opinion, other high level analytical models or they are estimated using empirical and physical models. An overview of axiomatic coverage models is presented in section 2. If statistical estimation is employed, then empirical models derive parameter estimates, including fault coverage, from data collected from physical models. The process used for parameter estimation is shown in Figure 2. Such empirical approaches are discussed in section 3. The parameter values are obtained via *fault injection* [3], [5], [15], [33] performed on physical models, which are discussed in Section 4.

2. Axiomatic Models of Fault Coverage

Axiomatic modeling of fault coverage is a behavioral representation of a system's response to faults. These models are embedded in the overall system model, and the actual number of coverage models required is a function of the system under test. There have been numerous refinements to the axiomatic fault coverage models and the various models that have been developed are presented in the following sections. These models are categorized into two sections: error handling without time limitation and error handling with time limitations.

2.1 Error Handling Without Time Limitations

The initial iteration of fault coverage models ignores any type of interference that could occur

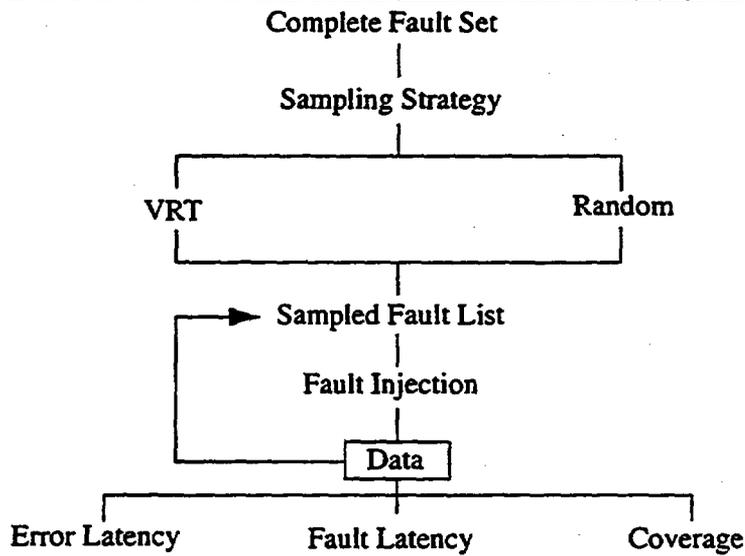


Figure 2. Parameter estimation hierarchy

during error handling, and typically consist of various forms of Markov and semi-Markov models. In these models, it is assumed that the time spent in states handling errors is negligible with respect to the time spent in states where errors are not present.

2.1.1 Permanent Effective Error Model [16]

The model, shown in Figure 3, depicts the effect of a fault and its resulting error. The fault

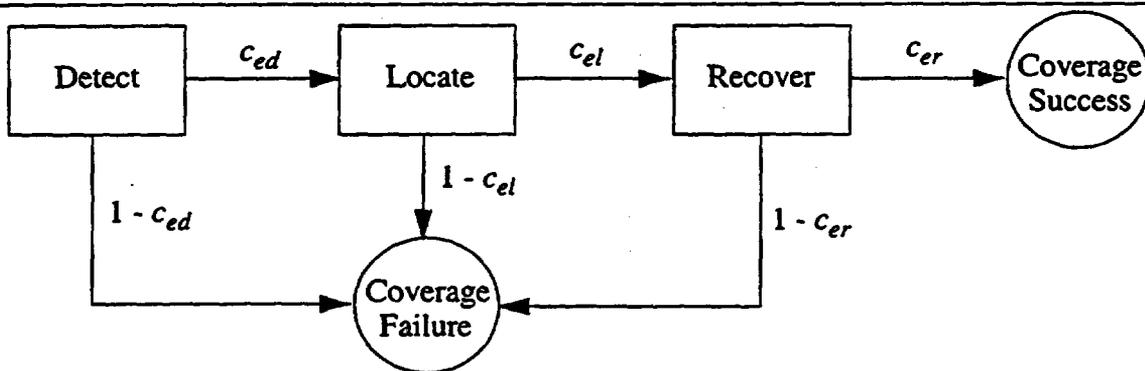


Figure 3. Permanent effective error model [16]

coverage for the system is given by

$$C = c_{ed} \times c_{el} \times c_{er} \quad (2)$$

where c_{ed} is the error detection probability, c_{el} is the error location probability, and c_{er} is the error recovery probability. Since this model only handles permanent faults and ignores transient faults, it has very limited applicability to real systems.

2.1.2 CAST Fault Coverage Model [8]

CAST, shown in Figure 4, combines transient fault restoration and permanent fault recovery. Faults occur at a rate $\lambda + \tau$, which is the sum of the permanent and the transient fault rates respectively.

Once a fault occurs, the detection state is entered with an error detection probability of c_{ed} . If the errors are not detected, then system failure occurs. However if the errors are detected, then transient recovery is attempted. The transient recovery probability is $1 - l$, where l is the transient leakage. If the transient recovery fails, then permanent recovery is attempted. In permanent recovery, the fault cause is located with probability c_{fl} and the system recovers with probability c_{sr} . If permanent recovery is successful, then $N-1$ modules remain. If it is unsuccessful, then system failure occurs. The n subscript associated with all of the system parameters simply denotes the number of active components.

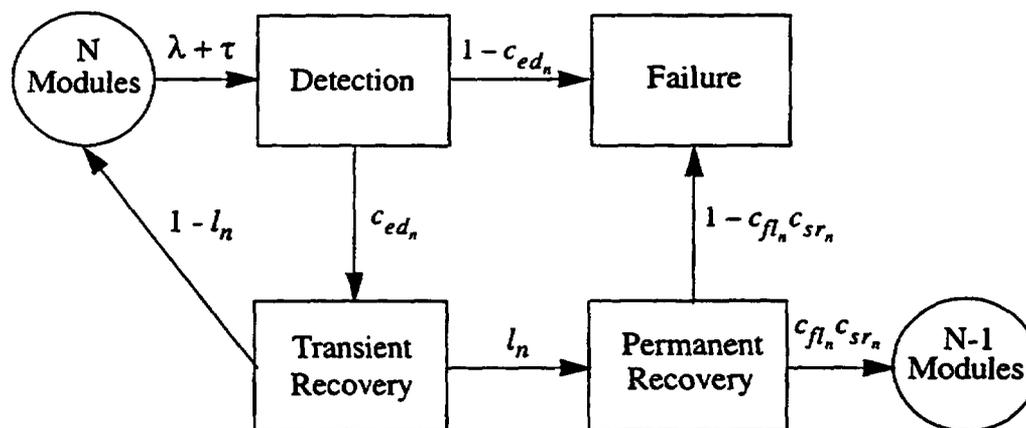


Figure 4. CAST fault coverage model [16]

2.1.3 CARE III Fault Coverage Model [36]

The CARE III single-error model, shown in Figure 5, is a generalized fault model realizing intermittent or permanent faults. In this model, state A represents the activation of an error. State B represents the error latency, where α_β and β_α are the transition rates between states A and B . State P represents the effects of the error polluting the system and occurs from state A at rate ρ . State D represents error detection, which can only occur if the error is active (state A) or it is polluting the system (state P). The rate at which an active error is detected before it can become latent or pollute the system is Δ , and the rate at which an error that is polluting the system becomes detected is $c_{ed}e_d$. If the error that is polluting the system is not detected, the error results in a failure, which is state F , at a rate of $(1 - c_{ed})e_d$ from state P . The probability of exit from state A to State D is given by

$$C = \frac{\Delta}{\Delta + \rho} + \frac{c_{ed}\rho}{\Delta + \rho} = \frac{\Delta + c_{ed}\rho}{\Delta + \rho} \quad (3)$$

In order to model permanent errors, α_β and β_α must be set to zero, which is the rate at which an effective error goes latent and vice versa; else, this model represents intermittent errors.

2.2 Error Handling With Time Limitations

In order for coverage models to be robust, consideration must be given to the lifetime of the fault and/or error. If the transient lifetime is considered, which in reality is a major concern, the models described in section 2.1 have very little applicability in developing accurate fault coverage estimation. The following models consider transient lifetime.

2.2.1 ARIES Fault Coverage Model [25]

The ARIES model, shown in Figure 6, includes permanent, transient and intermittent faults. In this model, there are three possible exits: (1) system crash; (2) normal processing; and (3) permanent fault

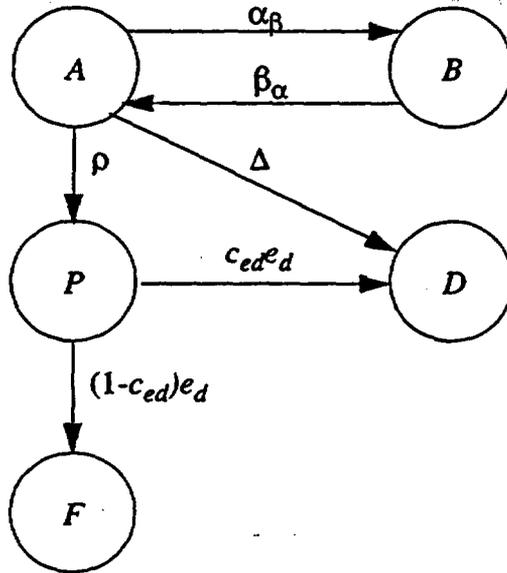


Figure 5. CARE III single-error fault model [16]

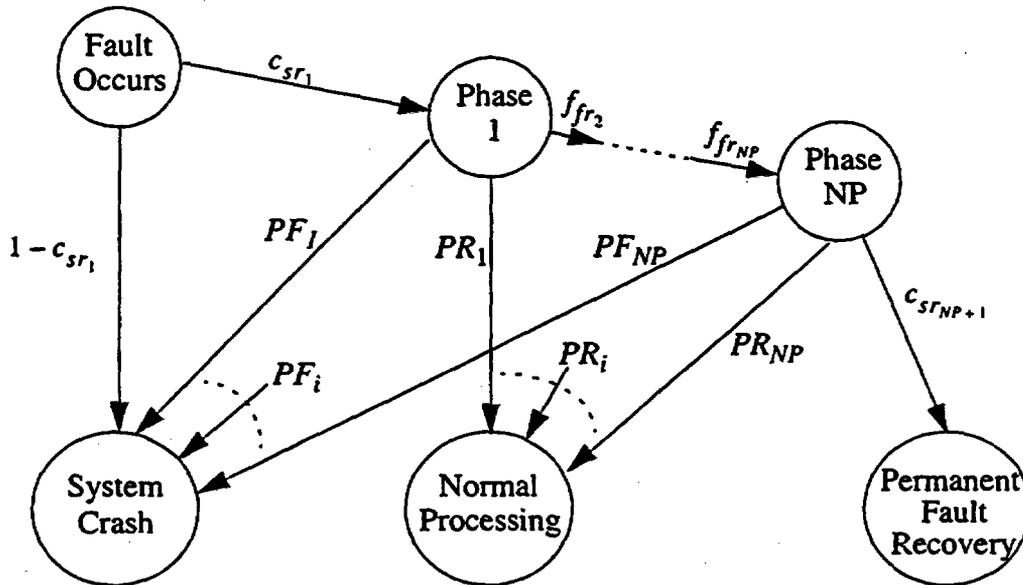


Figure 6. ARIES fault model [16]

recovery. Obviously, the system crash exit occurs when the error introduced by a fault causes system failure. The probability of a fault resulting in immediate system failure is $1 - c_{sr1}$. The fault recognition and attempted recovery probability is f_{fr_i} , where i denotes the recovery phase. The number of allowable recovery phases is fixed. If during a given recovery phase the system fails, then the system crash exit is taken with probability PF_i . If during a recovery phase the system recovers from a transient error, then the normal processing exit is taken with probability PR_i . Finally, if all recovery phases are entered and successful, then the permanent fault recovery exit is taken.

The ARIES fault coverage probability is

$$\begin{aligned}
 C &= \text{transient restoration probability} + \text{permanent error recovery} \\
 &= \sum_{i=1}^{NP} PR_i + c_{sr_{NP+1}} \times (\text{Coverage of Permanent Recovery Procedure})
 \end{aligned} \tag{4}$$

If the transient lifetime is long, then successful recovery may not have a fixed probability and the expression for PR_i must be modified.

If the transient lifetimes are exponentially distributed random variables (ν) and the duration of each recovery phase is a constant, the expression for transient restoration becomes

$$\begin{aligned}
 PR_i &= Pr[\text{Phase } i \text{ entered}] \times Pr[\text{Phase } i \text{ successful}] \\
 &\quad \times Pr[\text{transient gone before phase } i \text{ begins}] \\
 &= c_{sr_{i-1}} \times ER_i \times (1 - \exp(-(T_1 + T_2 + \dots + T_{i-1})/D))
 \end{aligned} \tag{5}$$

where the transient error lifetimes have a mean D with various durations. T_i and ER_i is the probability of an effective recovery procedure for phase i . This expression can be generalized for non-exponentially distributed error lifetimes as

$$PR_i = c_{sr_{i-1}} \times ER_i \times F_D(T_1 + T_2 + \dots + T_{i-1}) \tag{6}$$

where F_D represents the generalized distribution. If required, this type of generalization can be applied to the recovery phase.

2.2.2 Modified CARE III Model [36]

In [36], a transient error model is discussed in which transient lifetimes are assumed to be exponentially distributed. The duration of each recovery phase is assumed to be independent and identically distributed (iid), which is more restrictive than the ARIES model. However, it allows a random number of recovery phases, and like ARIES, it accommodates general distributions for the recovery phase. The CARE III model has been further refined to include transient, intermittent or permanent faults (errors), the effect of transient lifetimes and it can be solved for both the Markov and the semi-Markov case.

2.2.3 Extended Stochastic Petri Net (ESPN) Fault Coverage Model[16], [39]

The ESPN model combines both local and global timing. This model includes the limited recovery time associated with real systems, and the ability to determine the effects of near-coincident errors [26] that can occur during attempted recovery. The only near-coincident errors of interest are those whose occurrence can interfere with the current recovery. It is conservatively assumed that the occurrence of a near-coincident fault will result in system failure.

Since a stochastic Petri net is used, the various fault distributions can be generalized. If all faults are exponentially distributed ν , then the Petri net can be converted to a Markov model and solved accordingly. If the failure rates are not exponentially distributed, then in some cases the resulting model is semi-Markov and in other cases simulation is required.

2.3 Limitations of Axiomatic Coverage Models

As the development of axiomatic coverage models evolved, their ability to accurately model complex failure recovery mechanisms, such as the duration of an error with consideration given to its lifetime, expanded. In all of these models, however, there is one common thread: the model transition probabilities are unknowns. But these models are useful in determining which parameters are important, so that those gathering data know what data to collect.

It is impossible to know without actually testing a system the values of the various transition probabilities. In some circumstances, it may be impossible to ascertain the exact values. If feasible, a series of fault injection experiments can be performed on a physical model to try to obtain estimates for some of

these transition probabilities. Since testing the system for all possible faults is intractable, some type of sampling of the complete fault set for a system is required. Unless some empirical analysis is performed, such an estimation cannot be made. *Expert opinion* can be used to generate an abridged fault set, but it is impossible to demonstrate that such a fault set is complete to guarantee an accurate coverage estimate. Similarly, *expert opinion* can be used to simply predict the various model recovery and failure rates and the accuracy of such predictions is highly subjective. Additionally, each fault detection and recovery mechanism that resides in the system can require its own fault coverage model. Hence, the size and the complexity of a given axiomatic model increases relative to the size and the complexity of the system under test.

3. Empirical Models for Fault Coverage Parameter Estimation

The use of empirical models for fault coverage estimation requires detailed statistical analysis that must address four important questions [32]:

- (1) How can the fault coverage value be accurately estimated?
- (2) How can any error in the estimate be quantified?
- (3) How are fault samples selected?
- (4) How can accurate estimates for fault coverage be obtained in a reasonable time?

As previously discussed, empirical models are used to estimate parameters used by axiomatic model. Empirical modeling relaxes many of the assumptions and restrictions, such as parameter estimation, present in axiomatic models. Parameter estimation requires that the system fault space be sampled in some random fashion to provide a representative sample of the entire fault set. Using the data collected from this sampled set, statistical analysis is performed to analyze the accuracy of the resulting estimated parameters. It is shown in [4] that this technique can be used for predicting the system's expected fault coverage. There are numerous sampling strategies available, including techniques that attempt to reduce the variance of the estimate. This type of sampling is referred to as a variance reduction technique (VRT).

The purpose of VRTs is to increase the accuracy of the parameter estimate so that the required number of sample points can be further reduced. VRTs exploit some attribute of the system to increase the accuracy of the parameter estimate(s). Importance sampling, multi-stage sampling, stratified sampling and regression analysis are all examples of VRTs [9], [17].

3.1 Fault Coverage [12], [31], [32], [34], [35], [41]

The mathematical model used to describe a fault processing event is

$$\begin{aligned}
 C &= P(\text{fault processed correctly} | \text{fault existence}) \\
 &= P_{Y|X}(y|x) \\
 &= \frac{P_{X,Y}(x,y)}{P_X(x)}
 \end{aligned} \tag{7}$$

where $X \equiv \text{faults existence}$, and $Y \equiv \text{fault processed correctly}$. Since coverage is based upon a series of fault occurrences, that is a fault existence and its subsequent correct processing, the conditional probability in (7) can be considered to represent a series of discrete events. Hence, the expected value of the conditional probability for coverage can be modeled as

$$\begin{aligned}
 E[Y|X=x] &= \sum_{x,y \in \Omega} y \cdot P(Y=y|X=x) \\
 &= \sum_{x,y \in \Omega} y \cdot P_{Y|X}(y|x)
 \end{aligned} \tag{8}$$

where Ω is the system's complete fault space. Typically, a fault processing event is considered as a

Bernoulli ν defined as

$$y = \begin{cases} 1 & \forall \text{ covered faults with probability } p_{Y|X}(y|x) \\ 0 & \forall \text{ uncovered faults with probability } 1 - p_{Y|X}(y|x) \end{cases} \quad (9)$$

During testing, there is a possibility of *no-reply*, which is the inability to obtain measures from some elements in a sample [9]. Such problems arise when certain faults remain *hidden* when introduced to a system or it may be impossible to introduce a specific fault. In the preceding model, no-reply faults are not included. To remove this source of possible statistical error, the indicator function given in (9) is redefined as:

$$y_i = c(x_i) = \begin{cases} 1 & \forall \text{ covered faults} \\ 0 & \forall \text{ uncovered faults} \\ \partial & \forall \text{ no-reply} \end{cases} \quad (10)$$

and the analysis either counts the faults as covered, uncovered or discards the experiment. Substituting the expression for y found in (9) into (8) yields

$$\begin{aligned} E[Y|X=x] &= \sum_{x,y \in \Omega} y \cdot p_{Y|X}(y|x) \\ &= p_{Y|X}(y|x) \\ &= C \end{aligned} \quad (11)$$

Similarly, the variance of the conditional probability for coverage is

$$\begin{aligned} \text{Var}[Y|X=x] &= E[Y^2|X=x] - (E[Y|X=x])^2 \\ &= C(1-C) \end{aligned} \quad (12)$$

Obviously, neither the *pmf* associated with a fault's existence nor the joint *pmf* associated with a fault's existence and recovery is known *a priori*. As a result, a fault coverage experiment is necessary to determine a coverage value.

Theoretically, coverage can be determined by injecting the entire sample of N faults, which are assumed to be independent, from the fault space into a given system and calculating the ratio of properly handled faults, d , against the number of injected faults; that is [15],

$$C = \frac{d}{N} \quad (13)$$

The expression d , is analogous to that found in (9); that is, the number of properly handled faults in a fault injection experiment can be modeled as a summation of a series of Bernoulli trials. Hence,

$$d = \sum_{i=1}^N y_i \quad (14)$$

where y_i is the Bernoulli ν as defined in (9) for the i^{th} fault injection experiment and (13) can be rewritten as

$$C = \frac{1}{N} \sum_{i=1}^N y_i \quad (15)$$

Since it is impractical to inject every fault into a system, an experiment must be developed to provide an unbiased estimate of coverage by limiting the number of fault injection experiments.

By limiting the number of fault injection experiments to n , the coverage point estimate is

$$\hat{C}_{y_n} = \frac{1}{n} \sum_{i=1}^n y_i \quad (16)$$

which is an unbiased estimator if the equally likely constraint is valid [17]. Assuming that a large sample size exists, the central limit theorem (CLT) can be applied to approximate the estimator. Assuming that the estimator is Gaussian, it can be shown that its variance is

$$\text{Var}\{\hat{C}_{y_n}\} = \sigma_{C_{y_n}}^2 = \frac{\sigma_y^2}{n} \quad (17)$$

Since the fault coverage and the variance of y are unknown, the point estimate \hat{C}_{y_n} and the variance estimate $\hat{\sigma}_{C_{y_n}}^2$ are used in (17) [36] to yield

$$\hat{\sigma}_{C_{y_n}}^2 = \frac{\hat{\sigma}_y^2}{n} = \frac{\hat{C}_{y_n}(1 - \hat{C}_{y_n})}{n} \quad (18)$$

Under these conditions, a two-sided 100 γ % confidence interval can be defined. The lower bound of the confidence interval, which is the most conservative estimate of fault coverage, is of most interest and is defined in [29] as

$$\hat{C}_{y_{n,low}} = \hat{C}_{y_n} - \zeta_u \sqrt{\frac{\hat{\sigma}_{C_{y_n}}^2}{n}} = \hat{C}_{y_n} - \zeta_u \sqrt{\frac{\hat{C}_{y_n}(1 - \hat{C}_{y_n})}{n}} \quad (19)$$

where ζ_u is the confidence coefficient for a Gaussian distribution.

This statistical approach is the basis for many empirical models, which are reviewed in subsequent sections. In these models, VRTs are applied to provide variance reduction via various sampling techniques.

3.2 Powell et al Empirical Models [31], [32]

Since exhaustive testing to determine coverage is seldom possible, fault coverage estimation is performed on a representative sample of the entire fault space. There are two approaches for performing this random sampling: sampling from the complete fault space, and sampling from subspace partitions/classes of the complete fault list, which is commonly referred to as stratified sampling.

3.2.1 Non-partitioned Space Sampling

Representative sampling [24] consists of sampling with replacement from a group of n faults and is applicable to non-partitioned sampling. Its unbiased coverage estimator and variance are

$$\hat{C}_y = \frac{1}{n} \sum_{i=1}^n y_i \frac{p_{X,Y}(x_i, y_i)}{p_X(x_i)} \quad (20)$$

$$\text{Var}\{\hat{C}_y\} = \frac{1}{n} \left(\sum_{x_i \in \Omega} \left[y_i \frac{p_{X,Y}^2(x_i, y_i)}{p_X(x_i)} \right] - C_y^2 \right) \quad (21)$$

If the sample selection is chosen such that $p_X(x_i) = p_{X,Y}^2(x_i, y_i)$, then the estimate for the mean found in (20) is equivalent to the point estimate found in (16). Similarly, the variance for this estimate is given by (17) and the lower side of the confidence interval is given by (19).

3.2.2 Partitioned Space Sampling (Stratified Sampling)

Rather than sampling from the entire fault list, the sampling can occur from partitioned classes [9], [17], [24]. By definition, the classes form M disjoint subsets such that

$$E = \bigcup_{j=1}^M E_j \text{ such that for every } i, j, i \neq j, E_i \cap E_j = \emptyset \quad (22)$$

where E is the entire fault list. The coverage factor as expressed in (8) can be rewritten as

$$\begin{aligned} C &= \sum_{j=1}^M \sum_{x, y \in E_j} y \cdot p_{Y|X}(y|x) \\ &= \sum_{j=1}^M \sum_{x, y \in E_j} y \cdot p_{Y|X_j}(y|x_j) \cdot p_{X_j|X}(x_j|x) \\ &= \sum_{j=1}^M p_{X_j|X}(x_j|x) \sum_{x, y \in E_j} y \cdot p_{Y|X_j}(y|x_j) \\ &= \sum_{j=1}^M p_{X_j|X}(x_j|x) c(X_j) \text{ where } c(X_j) = \sum_{x, y \in E_j} y \cdot p_{Y|X_j}(y|x_j) \end{aligned} \quad (23)$$

Using this partitioned sampling space, two different sampling techniques can be implemented: the naive estimator and stratified sampling.

The *naive estimator* samples an equal number of faults from each class. For each sample, the coverage estimate for each class i is

$$\hat{C}_{na} = \frac{1}{n} \sum_{i=1}^M d_i = \frac{d}{n} \quad (24)$$

The estimator's variance is

$$\text{Var}\{\hat{C}_{na}\} = \frac{1}{nM} \sum_{j=1}^M (c(X_j) - \hat{c}(X_j))^2 \quad (25)$$

If all fault occurrences are not equally probable, then this estimator is biased. It can be shown that

$$E\{\hat{C}_{na}\} = \hat{c}(X_j) = \frac{1}{M} \sum_{j=1}^M c(X_j) \quad (26)$$

hence this technique provides a *naive estimator*.

The covariance between the coverage, C_{E_j} , and the fault occurrence probability, $p_X(X_j)$, for each class is

$$S_{CP} = \frac{1}{M} \sum_{j=1}^M (c(X_j) - \hat{c}(X_j)) \cdot \left(p_{X_j|X}(x_j|x) - \frac{1}{M} \right) \quad (27)$$

from which it can be proven that

$$\hat{c}(X_j) = C - MS_{CP} \quad (28)$$

Depending on the sign of the covariance, the fault coverage estimator can be either pessimistic or

optimistic. This result is proven with actual examples in [31], [32].

In *stratified sampling* [9],[17], [24], a number of samples, n_j , for each class, E_j , is pre-selected and a representative sample of size $n_i = n_j$ is taken for each class. The coverage factor now applies to the class rather than the complete sample space and is expressed as

$$\hat{C}_{\Omega_i} = \frac{d_i}{n_i} \quad (29)$$

The coverage and variance estimates are

$$\hat{C}_{stE} = \sum_{i=1}^M p_{X_i|X}(x_i|x) \cdot \hat{c}(X_i) \quad (30)$$

$$Var\{\hat{C}_{stE}\} = \sum_{i=1}^M p_{X_i|X}(x_i|x) Var\{\hat{c}(X_i)\} \quad (31)$$

Similarly, the variance of the class coverage estimator is given by

$$Var\{\hat{c}(X_i)\} = \frac{1}{n_i - 1} (\hat{c}(X_i) - \hat{c}^2(X_i)) \quad (32)$$

From these variance expressions, it can be seen that the variance depends upon the class sample size. To minimize the system coverage factor's variance, $Var\{\hat{C}_{st\Omega}\}$, each class' sample size must be defined as

$$n_i = p_{X_i|X}(x_i|x)n \quad (33)$$

This type of sample size allocation is referred to as a *stratified sample with representative allocation*. Using the expressions found in (33), (29) and (30), the system coverage estimator can be expressed as

$$\hat{C}_{stR} = \frac{d}{n} \quad (34)$$

which is equivalent to that for the *naive estimator*. The variance, however, differs. If the expression n_j is substituted into (31) and (32), the resulting variance is

$$Var\{\hat{C}_{stR}\} = \frac{1}{n} C - \frac{1}{n} \sum_{i=1}^M p_{X_i|X}(x_i|x) c^2(X_i) \quad (35)$$

Hence, the precision of representative stratification is not sensitive to the covariance between the coverage and the fault/activity occurrence probability for each class. As a result, there is an appreciable gain in precision for coverage estimation and this is substantiated with examples from [31], [32]. This gain in precision is demonstrated via the improvement in the confidence interval obtained by using the variance provided by (35) in (19). However, the fault/activity occurrence probability is an unknown, and as a result, it is difficult for representative stratification to be used accurately.

3.2.3 No-Reply Problem

To accommodate the no-reply problem, *a posteriori* stratification is introduced. This method uses available system information in considering different stratification techniques. Since structural information is circuit dependent, the selected stratification technique is viable only for the circuit under test. Hence, this methodology cannot be extended to a general application.

3.3 Cukier et al Model [14]

The Cukier et al model is an extension of the work performed by Powell et al. In this work, fault

coverage is modeled in terms of the uncovered faults. This non-coverage estimate, \hat{C} , using representative sampling is

$$\hat{C} = \frac{\bar{d}}{N} \quad (36)$$

where \bar{d} are the number of uncovered faults and N is the number of fault injection experiments performed. The expression for non-coverage shown in (36) is analogous to the expression for coverage shown in (13). Similarly, the coverage expression for partitioned space sampling, which is shown in (23), can be expressed in terms of non-coverage as

$$\bar{C} = \sum_{j=1}^M p_{X_j|X}(x_j|x) \bar{c}(X_j) \quad (37)$$

Two distinct approaches for non-coverage estimation can be made using either classical statistical methods or Bayesian techniques.

3.3.1 Classical Statistical Approach

The upper 100 γ % confidence limit estimator for \bar{C} is defined as

$$P[\bar{C} \leq \bar{C}_\gamma(X) | \bar{C}] = \gamma \quad (38)$$

In modeling non-coverage for an ultra-dependable system, it is shown in [30] that approximated estimations using the classical statistical approach are not valid. Hence, approximations cannot be used when developing non-coverage estimators based upon (38).

To allow for the multiple classes during the fault injection experiments and to minimize (37), the upper 100 γ % confidence limit estimator for \bar{C} for M classes is given by the solution of

$$\prod_{i=1}^M \left(1 - \sum_{x_i=0}^{x_i} \binom{n_i}{x_i} \bar{c}_i^{x_i} (1 - \bar{c}_i)^{n_i - x_i} \right) \forall i \in \{1, \dots, M\}, \bar{c}_i \in [0, 1] \quad (39)$$

3.3.2 Bayesian Approach

In Bayesian theory, non-coverage, \bar{C} , and the class non-coverages, \bar{c}_i , are considered as ν . The upper 100 γ % confidence limit is defined by the distribution of the ν ; that is

$$P[\bar{C} \leq \bar{C}_\gamma(X) | X] = \gamma = x \quad (40)$$

In order to obtain the confidence limit defined in (40), the posterior distribution of \bar{C} is required. For representative sampling, this posterior distribution is

$$f_{\bar{C}}(\bar{c} | X = x) \quad (41)$$

and for stratified sampling, the posterior distribution of the non-coverage classes is simply

$$f_{\bar{C}_i}(\bar{c}_i | X_i = x_i) \quad (42)$$

In order to solve for the posterior distribution, an appropriate choice of the prior distribution for the non-coverage classes is required.

A beta prior distribution is used for two primary reasons: (1) the number of uncovered faults in each partition is binomially distributed and a beta prior distribution ensures that the prior and the posterior distributions are both from the same family; and (2) when the parameters of the beta distribution equal one, the obtained distribution is uniform over the interval [0, 1], which means that all values of \bar{c}_i have the same weight. The beta prior distributions for \bar{C}_i are

$$f_{\bar{C}_i}(\bar{c}_i) = \frac{\bar{c}_i^{-k_i-1} (1-\bar{c}_i)^{l_i-1}}{\beta(k_i, l_i)} \text{ for } 0 \leq \bar{c}_i \leq 1, k_i > 0, l_i > 0 \quad (43)$$

where $\beta(k_i, l_i)$ is a beta function with parameters k_i and l_i [1].

Since the number of uncovered faults in each partition, X_i , is binomially distributed, then

$$f_{X_i}(x_i | \bar{C}_i = \bar{c}_i) = \binom{n_i}{x_i} \bar{c}_i^{x_i} (1-\bar{c}_i)^{n_i-x_i} \quad (44)$$

and the posterior distribution for \bar{C}_i is [2]

$$f_{\bar{C}_i}(\bar{c}_i | X_i = x_i) = \frac{\bar{c}_i^{-k_i-1} (1-\bar{c}_i)^{l_i-1}}{\beta(k_i', l_i')} \quad (45)$$

where $k_i' = x_i + k_i$ and $l_i' = n_i - x_i + l_i$.

The posterior distribution for \bar{C} is found by combining the posterior distributions of the various \bar{C}_i . In [13], it is shown that an analytical expression for the posterior of \bar{C} is too complicated for more than three classes. Thus, the posterior distribution for \bar{C} can only be obtained using approximation. When a distribution cannot be exactly calculated, it is possible to theoretically exhibit all of the properties of a distribution in terms of its moments [37]. Similarly, distributions that have a finite number of lower order moments in common approximate each other [37]. The calculation of moments can be achieved using either the moment generating function [29], [37] or assuming independence among the classes. Once the moments have been calculated, the posterior distribution can be determined from the Pearson distribution system [21].

3.3.2.1 Calculation of Moments

The moment generating function of \bar{C}_i , assuming a Beta distribution $\beta(k_i', l_i')$, is

$$\phi_{\bar{C}_i} = -F(k_i'; k_i' + l_i'; t) \quad (46)$$

where $-F(k_i'; k_i' + l_i'; t)$ is the confluent hypergeometric function [48] and

$$\phi_{p_i \bar{C}_i} = -F(k_i'; k_i' + l_i'; p_i t) \quad (47)$$

Since the moment generating function of a sum of rv is equivalent to the product of the moment generating function of the various rv [20], then

$$\phi_{\bar{C}}(t) = \prod_{i=1}^M -F(k_i'; k_i' + l_i'; p_i t) \quad (48)$$

which is derived based upon (23). The n^{th} derivative of the moment generating function of \bar{C} for $t=0$ defines the n^{th} moment of \bar{C} . Assuming that the powers of \bar{C}_i are independent, then simpler expressions for the moments of \bar{C} can be obtained. The r -th central moment of the beta distribution, $\beta(k_i', l_i')$, using this independence assumption is

$$E[(X - \mu_X)^r] = \frac{\beta(k_i' + r, l_i')}{\beta(k_i', l_i')} = \frac{k_i'(k_i' + 1) \dots (k_i' + r - 1)}{(k_i' + l_i')(k_i' + l_i' + 1) \dots (k_i' + l_i' + r - 1)} \quad (49)$$

3.3.2.2 Pearson Distribution System [21] for Use as a Posterior Distribution

The Pearson distribution system is a family of seven distributions and are summarized in Table 1. The seven distributions are represented in a planar plot of their skewness and kurtosis coefficients. From this planar plot, the family to which a given data set belongs is determined. The Pearson distribution pdf,

Table 1: Pearson Distribution System [21]

| | | | |
|----------|---|----------|---|
| Type I | beta distribution | Type V | inverse Gaussian distribution |
| Type II | symmetrical form of the function defined in Type I | Type VI | cumulative Pareto distribution representing <i>income</i> |
| Type III | gamma distribution | Type VII | <i>t</i> distribution |
| Type IV | no common statistical distributions are of this type; the values required for the CDF are intractable | | |

$f(x)$, satisfies a differential equation of the form

$$\frac{1}{f} \frac{df}{dx} = -\frac{pa+x}{pb_0+pb_1x+pb_2x^2} \quad (50)$$

The shape of the distribution is dependent on the values of the four parameters, which can be determined by the first four moments of the distribution $f(x)$. For a detailed summary of this relationship, the reader is advised to see [1], [2], [13],[14], [21],[30], [37].

3.3.3 Comparison of Approaches

The classical statistical and the Bayesian approach are compared for two hypothetical systems, system 1 and system 2 [14], using stratification and simple sampling. It is assumed that the prior distribution for the Bayesian estimation is uniform; that is, the parameters of the beta distribution are equal to one. Initially, the moments used for the Bayesian analysis are calculated using both moment generating functions and the independence assumption.

The initial testing uses homogenous allocation, which requires sampling a predetermined number from each class, and representative allocation, which requires sampling the same number of faults from each class. During this testing the number of fault injection experiments that are performed is varied to determine the validity of the Bayesian approach and to compare it to the classical statistical approach.

During testing, it is shown that only the moment generating function when used with representative allocation produces valid results for system 1; that is, the posterior distribution is of Type I. Both estimation methods are valid for system 2 when used with representative sampling. Hence, the comparison is performed using the Bayesian method is derived via moment generating functions and using representative sampling. When simple sampling is considered, the Bayesian estimations are more conservative. Using stratification, it is shown that the Bayesian estimation is less conservative than the classical statistical methods. However, this conservatism decreases as the number of fault injection experiments increases.

3.4 Fault Expansion Model [34], [35]

Another method for sampling the complete fault space is fault expansion. In fault expansion, the fault space is subdivided into mutually exclusive subsets defined as

$$E_i = \left\{ x_{i1}, x_{i2}, \dots, x_{i|E_i|} \right\} \quad (51)$$

where E_i is the i -th equivalent fault set, x_{ij} is the j -th element of the i -th fault set and $|E_i|$ is the set cardinality. All equivalent fault sets are disjoint and their union is the complete fault set.

The fault expansion process consists of randomly selecting a fault and determining the set of equivalent faults. All members of E_i are removed from the fault space and fault injection is performed

using only one fault. The evaluation of the x_{ik} fault is described mathematically as

$$z_i = c(x_{ij}) = \begin{cases} 1 \forall \text{covered faults} \\ 0 \forall \text{uncovered faults} \end{cases} \quad (52)$$

where z_i is the i^{th} sample of the z rv which describes the result of the fault injection experiment for the equivalent fault set E_i . The expected value for coverage is

$$E\{Z|X\} = \sum_{x, z \in E} z \cdot P_{Z|X}(z|x) = \sum_{i=1}^{|E|} \sum_{j=1}^{|E_i|} z_{ij} P_{Z|X}(z_{ij}|x_i) = C \quad (53)$$

which is similar to (8). There have been two VRTs developed using fault expansion and they are the Wang et al empirical model [41] and the Smith et al empirical model [35].

3.4.1 Wang et al Empirical Model [30]

If fault sampling occurs for the entire fault space, the total number of covered faults after m injections is simply

$$C_m = \sum_{i=1}^m X_i \quad (54)$$

Using the binomial distribution for C_m , the 100 γ % one-sided confidence interval [9], [38] for the coverage estimate is

$$P(C_m \geq c_m | dc_c) = \sum_{j=c_m}^m \binom{m}{j} dc_c^j (1 - dc_c)^{m-j} = 1 - \gamma \quad (55)$$

where γ is the confidence coefficient and dc_c is the desired coverage value. It is very difficult to solve (55) for dc_c given an arbitrary value of m .

For a system with coverage near one, a Poisson distribution is a good approximation to the binomial distribution. In this case, it can be shown that dc_c is given by [38]

$$dc_c = 1 - \frac{1}{2m} \chi_{deg; 1-\gamma}^2 \quad (56)$$

where $\chi_{deg; 1-\gamma}^2$ satisfies $P(Y > \chi_{deg; 1-\gamma}^2) = 1 - \gamma$ and Y is chi-square distributed with $deg = 2(k - s_m + 1)$ degrees of freedom. In testing, it is determined that for coverages approaching one, the value of c_m is extremely close if not equal to m . To ensure that the lower limit for the confidence interval is met or exceeded, the value of m must be extremely large. To reduce the required sample size and to meet the lower confidence interval requirement, fault expansion [34], [35] is used.

In sampling using fault expansion, there are two cases of interest: the infinite and the finite fault population. For the infinite population, it is shown in [34] that the best estimate for fault coverage occurs when all fault classes are of equal size. The resulting lower one-sided confidence interval for this coverage estimate is identical to that found in (55). Since there is no appreciable variance reduction, fault expansion is not recommended. However, fault expansion is very helpful for the finite population case.

Assuming that the fault population is finite the exact coverage factor is given by (13). Since it is impractical to inject all N faults, the value for D must be estimated. It is proven that the one-sided confidence interval for the lower limit on the estimate of D , that is D_l , using the binomial distribution is

$$P(C_{fi} \geq c_{fi} | DF_l) = \frac{\Gamma\left(dc_l \frac{N}{b} + 1\right) \Gamma\left(\frac{N}{b} - fi + 1\right)}{\Gamma\left(dc_l \frac{N}{b} - fi + 1\right) \left(\Gamma\left(\frac{N}{b} + 1\right)\right)} = 1 - \gamma \quad (57)$$

where b is the equivalence class size, N is the finite fault population size, f_i is the total number of injections, dc_l is the lower limit on coverage and γ is the confidence coefficient.

From (57), it can be determined that dc_l is a function of b/N and f_i . In analyzing this equation, it is proven that the size of the population classes greatly affects the analysis. If the equivalence classes are significantly large, the impact of fault expansion is maximized. This result implies that unlike in the infinite population case, the size of the fault classes need not be equivalent; instead, it is desirable that the equivalent fault classes for covered faults should be considerably larger than those for uncovered faults.

3.4.2 Smith et al Empirical Model [34], [35]

If all sample measurements are covered, then the variance found in (18) is zero. Hence, no confidence interval can be calculated. To overcome this limitation, a more conservative variance estimate is needed. By converting one covered fault injection experiment into an uncovered experiment, the variance estimate is always non-zero and more conservative in nature; that is, the calculated variance will exceed the actual value. This modified variance estimate of y is

$$\hat{\sigma}_{y_n}^2 = \frac{n-1}{n^2} \quad (58)$$

and the resulting confidence interval for the lower bound is

$$\hat{C}_{y_{n_{low}}} = 1 - \zeta_{nN} \sqrt{\frac{\hat{\sigma}_{y_n}^2}{n}} = 1 - \zeta_{nN} \sqrt{\frac{n-1}{n}} \quad (59)$$

where ζ_{μ} is selected depending upon the desired confidence level. This lower bound is consistent with other analyses of the all covered case [27].

The point estimate mean for \hat{C}_{z_n} is

$$\hat{C}_{z_n} = \frac{\sum_{i=1}^n z_i |E_i|}{\sum_{i=1}^n |E_i|} = \frac{1}{m} \sum_{i=1}^n z_i |E_i| \quad (60)$$

where

$$m = \sum_{i=1}^n |E_i| \quad (61)$$

Assuming that the sample size is sufficiently large, then the CLT can be applied, and the resulting point estimate for \hat{C}_{z_n} is

$$\hat{C}_{z_n} = \sum_{i=1}^n z_i p_i \quad (62)$$

where $p_i = |E_i|/m$, $\sum_{i=1}^n p_i = 1$ and p_i is the probability that a fault lies in class i . The resulting variance is

$$\hat{\sigma}_{z_n}^2 = \sum_{i=1}^n (z_i - \hat{C}_{z_n})^2 p_i \quad (63)$$

Similarly, the variance reduction is derived assuming there exists one uncovered fault set.

Assuming that the first class E_1 is uncovered, the uncovered probability p_1 is

$$p_1 = \frac{|E_1|}{m} \quad (64)$$

and the point estimate as described in (62) becomes

$$\hat{C}_{z_n} = z_1 p_1 + \sum_{i=2}^n z_i p_i = \frac{m - |E_1|}{m} \quad (65)$$

and the variance becomes

$$\hat{\sigma}_{z_n}^2 = \frac{|E_1|(m - |E_1|)}{m^2} \quad (66)$$

By assuming that the covered and the uncovered probabilities are of the following form

$$p_c = \frac{1}{m} \sum_{i \in \text{covered faults}} |E_i| \quad (67)$$

$$p_c = \frac{1}{m} \sum_{i \in \text{uncovered faults}} |E_i| = \frac{|E_c|}{m}$$

(66) can be rewritten as

$$\hat{\sigma}_{z_n}^2 = \frac{|E_d|(m - |E_d|)}{m^2} \quad (68)$$

As long as $|E_c| < (um)/n$, where \bar{d} is the number of uncovered fault injection experiments, the variance is reduced. If the average uncovered fault class size is smaller than the average set size, then fault expansion provides variance reduction.

As is true for the random sampling case, a conservative estimation must be made for the all covered case. It is again conservatively estimated that one of the covered fault injection experiments is assumed to be uncovered to prevent a zero variance. To minimize the increase in variance, $|E_d|$ in (68) is set equal to one. The resulting variance estimate is

$$\hat{\sigma}_{z_n}^2 = \frac{m-1}{m^2} \mathbb{V} \left\{ \left(\sum_{i=1}^n z_i \right) = n \right\} \quad (69)$$

The variance reduction ratio for \hat{C}_{z_n} as it relates to the original \hat{C}_{y_n} point estimate is

$$v_r = \frac{\hat{\sigma}_{z_n}^2}{\hat{\sigma}_{y_n}^2} = \frac{\frac{|E_c|(m - |E_c|)}{m^2}}{\frac{\left(\sum_{l=1}^n y_l \right) \left(n - \sum_{l=1}^n y_l \right)}{n^2}} = \frac{n^2(|E_c|(m - |E_c|))}{m^2 \left(\sum_{l=1}^n y_l \right) \left(n - \sum_{l=1}^n y_l \right)} \quad (70)$$

which is a measure of statistical improvement that results from utilizing fault expansion. For the all covered case, the resulting variance reduction ratio is

$$v_r = \frac{\hat{\sigma}_{z_n}^2}{\hat{\sigma}_{y_n}^2} = \frac{(m-1)n^2}{(n-1)m^2} \cong \frac{n}{m} \text{ if } (m \gg 1) \text{ and } (n \gg 1) \quad (71)$$

3.5 Constantinescu Empirical Model [12]

In this model, multi-stage, stratified and combined multi-stage/stratified random sampling [9] are used. Multi-stage sampling allows the use of a multidimensional event space. The number of dimensions is equivalent to the number of factors that affect coverage. This event space consists of the cross products of all possible fault locations, types, times of occurrence, durations and system workloads. Due to practical considerations, a 3-D space is used consisting of the cross products of three 1-D subspaces. These subspaces include fault location, fault occurrence times and system input values. The population is divided into consecutive *subunits*, and random sampling is performed on these newly created subunits. The coverage model found in (16) is re-written to accommodate a multi-dimensional sample space as

$$C = \frac{\sum_{i_1=1}^{|E_1|} \cdots \sum_{i_n=1}^{|E_n|} y(i_1, i_2, \dots, i_n)}{\prod_{i=1}^n |E_i|} \quad (72)$$

Obviously, the number of fault locations is finite by nature. Faults can occur at any time, but time can be subdivided into a finite number of small intervals. The input space, however, is not as easily defined. As system complexity increases, the number of input values becomes extremely large. To overcome this problem, stratification is used to manage the vast input space by subdividing it into smaller, more manageable subspaces called *strata*. The sum of the strata equals the original population. Each stratum is sampled *s*-independently. By combining both multi-stage and stratified sampling techniques, the effectiveness of sampling increases [9], [40].

3.5.1 Multi-Stage Sampling for a 3-D Space

Multi-stage sampling requires random selection of members from the original population followed by consecutive random sampling from the subunits. The three sampling stages for simulated fault injection and for physical fault injection performed in 3-D space are found in Table 2. Regardless of the sampling

Table 2: Fault Injection Sampling Stages

| | Simulated Fault Injection | Physical Fault Injection |
|---------|--|--|
| Stage 1 | Select several input values at random from the 1-D input space | Randomly select the locations for fault injection for all previously selected inputs and injection times |
| Stage 2 | Select several fault injection times at random from the 1-D fault occurrence time space for each input | Select several input values at random from the 1-D input space |
| Stage 3 | Randomly select the locations for fault injection for all previously selected inputs and injection times | Select several fault injection times at random from the 1-D fault occurrence time space for each input |

order that is used, the formulae for coverage and general multi-stage sampling are applicable. Unbiased estimates for the mean and variance [11] are

$$\hat{C} = \frac{\sum_{i_1=1}^{|w_1|} \cdots \sum_{i_n=1}^{|w_n|} y(i_1, i_2, \dots, i_n)}{n \prod_{j=1}^n |w_j|} \quad (73)$$

and

$$Var(\hat{C}) = \sum_{k=1}^n \left\{ \frac{\left[\prod_{k'=1}^{k-1} \left(1 - \frac{|w_{k'}|}{|E_{k'}|} \right) \right] |w_k|}{\prod_{k'=1}^k |w_{k'}|} \left(\frac{\sum_{p, i_1=1}^{|w_1|} \cdots \sum_{i_n=1}^{|w_n|} \left(y(i_1, i_2, \dots, i_n) - \sum_{i_n=1}^{|w_n|} y(i_1, i_2, \dots, i_n) \right)^2}{\prod_{i=1}^{n-1} |w_i| (|w_n| - 1)} \right) \right\} \quad (74)$$

where $|w_k|$ is the sample size and $|E_k|$ is the population size at stage k for n sampling stages. The confidence level is as defined in (19) and the indicator function is modeled in terms of a multi-dimensional space. Theorem proofs for these estimates can be found in [10].

3.5.2 Stratified Sampling

In many situations, the amount of system input data needed is too large to explicitly input all data combinations. By using stratification to divide the input space into strata, the maximum number of possible inputs is greatly reduced. Typically, the maximum number of inputs grouped in a stratum is determined by the largest random number that can be practically generated, and by the type of inputs, binary or analog. For the 3-D event space for coverage, the original event space, E , is subdivided into several smaller subspaces.

Assuming the point estimate is Gaussian, the unit population, E , be subdivided into m subpopulations. If independent random sampling is performed in every stratum, then the unbiased estimate of the mean is

$$\hat{C}_{ST} = \sum_{i=1}^m f_{E_i}(st) \hat{C}_i \quad (75)$$

$$p_{E_i}(st) \equiv \frac{E_{ST_i}}{E}$$

where ST stands for stratified. An unbiased variance estimate is

$$Var(\hat{C}_{ST}) = \sum_{i=1}^m p_{E_i}^2(st) \cdot Var(\hat{C}_i) \quad (76)$$

Theorem proofs for these estimates can be found in [10].

3.5.3 Combined Multi-Stage and Stratified Sampling

Coverage estimator equations for combined multi-stage and stratified sampling are obtained from the simpler expressions found in theorems 1-6 and corollary 1 in [12]. For example, the appropriate sampling order for simulated fault injection is stratification followed by multi-stage sampling. The weight

of stratum j is

$$f_{E_j}(st) = \frac{\prod_{i=1}^n E_{ST_{ji}}}{\sum_{j=1}^m \prod_{i=1}^n E_{ST_{ji}}} \text{ for } i = 1, 2, \dots, n \quad (77)$$

The resulting confidence limits on the coverage can be found by substituting (76) using the stratum weights derived in (77) into (19).

3.6 Limitations of Empirical Models for Coverage Estimation

Empirical models approach coverage estimation by implementing various VRTs. In the Powell et al method, the variance reduction is achieved using stratified sampling and two-stage sampling. However, the conditional probabilities that are required for this estimate need to be measured. These measurements can be extremely difficult, if not impossible, and the problems associated with estimating these probabilities are demonstrated by the wide range of coverage estimations shown in [31], [32]. In two-stage sampling, the results demonstrated that better coverage estimation can be achieved than for stratified sampling. However, for both of these methods, it is concluded that the coverage estimate is system dependent.

In the Wang et al model, fault expansion is used to provide variance reduction. In this model it is assumed that the binomial representation for coverage approximates a Poisson distribution, which is then used to calculate a single sided confidence level. In the model presented by Smith et al, they applied the CLT to determine the confidence interval associated with the coverage estimation, because it is assumed that the sample size is sufficiently large [29]. Similarly in the Constantinescu method, his point estimates for the various sampling techniques discussed assumes that the point estimates are Gaussian. As a result, these assumptions limit the broad applicability of these methods. If the underlying distribution for the point estimate is not Gaussian or Poisson, then none of these approximations are appropriate.

The coverage estimation model by Cukier et al is a refinement of the Powell et al model using uncovered fault information to develop a Bayesian estimate. This model requires the conditional probabilities from the Powell et al model, which are difficult if not impossible to measure, for calculating both the moments and the Pearson parameters in the Bayesian approach. Additionally, it is assumed that the prior distribution for the coverage estimate is beta with parameters one. If however the prior distribution is not beta or the beta parameters differ, then the calculation of the posterior distribution is in error. Finally, this model also shows some dependence on the system being modeled; that is, the ability to derive a coverage estimate can be system dependent.

4. Physical Models

Physical models represent the actual system and involve the development of prototypes realized in either software and/or hardware. Additionally, these models can be realized at multiple levels of abstraction such as the transistor, gate, circuit or system level, allowing for hierarchical modeling. Ideally, all parameters can be measured at these various levels, but there is no accepted way to keep the faults modeled consistently throughout the various levels of hierarchy. Nevertheless, the parameter estimates' measure should be consistent and should improve as the level of modeling descends to the lowest level, which is the transistor level.

The feasibility of constructing and testing hardware prototypes depends upon both the time available and the cost of production. If only time is a concern, then a sample system can be constructed and tested in a harsh (that is, failure rate accelerated) environment. If cost is a concern, then software based prototypes can be built. The obvious problem here is the time required to simulate the models.

These software models can be tested just like an actual sample system using fault injection. For both of these techniques, an unbiased subset of the overall fault space must be used for fault injection. As discussed in section 2.3, the selection of a fault list subset that is of sufficient size to provide an accurate

fault coverage estimate requires an empirical model. If this analysis is not performed, then the selection of the list of faults to be tested is highly subjective and it may not provide a statistically meaningful result.

5. Conclusions

Fault coverage can be examined in two distinct ways: (1) fault coverage modeling and (2) parameter estimation. Fault coverage modeling consists of creating an axiomatic model that represents faulty component behavior within a given system. One drawback of this approach is that the size and the complexity of a system model would dramatically increase because of the state space required to represent all possible fault scenarios. Additionally, the parameters that these axiomatic models require, such as the transition rates between the various states, are not known *a priori*. As a result, these values are approximated from other analytical models or expert opinion, and the resulting fault coverage is simply an approximation with undefined confidence intervals. To overcome these problems, empirical and physical models are used.

Empirical models are statistical models that use data collected from physical models. Using fault injection with physical models, data pertaining to the various dependability parameters, including fault coverage, is collected and parameter estimation can be made from the empirical models. Dependability testing must be incorporated during the design cycle, and via fault injection, the various dependability parameters, including fault coverage, can be estimated for a given confidence interval from the empirical models using VRTs.

Fault coverage estimation is typically achieved via point estimation and Bayesian techniques. Point estimation, assuming that a large enough data set exists, can use the CLT, which implies a Gaussian distribution, from which a two-sided confidence interval can be calculated. Special attention is given to the lower bound because it is the most pessimistic estimate. Additionally, a coverage point estimate can be obtained if coverage is assumed to be binomially distributed. This distribution can be estimated by a Poisson distribution from which a single-sided confidence interval can be extracted. In all of these methods, various sampling schemes are implemented to provide variance reduction for the coverage estimates.

The limitation of these approaches is that the empirical models base their parameter estimates upon an *a priori* selection for the distribution of the point estimate, and in the case of Bayesian techniques, the coverage estimate is based upon an *a priori* selection of the prior distribution. As discussed in section 3.6, this *a priori* selection of the various distributions limits the applicability of these existing empirical models. Additionally, some of these models are system dependent, which further limits their applicability.

References

- [1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Dover Publications, New York, 1972.
- [2] J. Aitchison and I. R. Dunsmore, *Statistical Prediction Analysis*, Cambridge University Press, Cambridge, England, 1975.
- [3] J. Arlat, M. Aguera, L. Amat, Y. Crouzet, J.-C. Fabre, J.-C. Laprie, E. Martins and D. Powell, "Fault Injection for Dependability Validation: A Methodology and Some Applications," *IEEE Tran. Software Engineering*, vol. 16, no. 2, Feb. 1990, pp. 166-181.
- [4] J. Arlat, A. Costes, Y. Crouzet, J.-C. Laprie, and D. Powell, "Fault Injection and Dependability Evaluation of Fault Tolerant Systems," *IEEE Trans. Computers*, vol. 42, no. 8, Aug. 1993, pp. 913-923.
- [5] J. Arlat, Y. Crouzet and J.-C. Laprie, "Fault Injection for the Experimental Validation of Fault Tolerance," *Proc. Esprit Conference*, Nov. 1991, pp. 791-805.
- [6] W.G. Bouricius, W.C. Carter and P R. Schneider, "Reliability Modeling Techniques for Self-repairing Computer Systems," *Proc. 24th ACM Conference*, Mar. 1969, pp. 295-309.
- [7] W. C. Carter and J. Abraham, "Design and Evaluation Tools for Fault Tolerant Systems," *Proc. AIAA Computers in Aerospace Conference*, 1987.
- [8] R.B. Conn, P.M. Merryman and K.L. Whitelaw, "CAST - A Complimentary Analytic-Simulative Technique for Modeling Fault Tolerant Computing Systems," *Proc. AIAA Computers and Aerospace Conference*, Nov. 1977

- [9] W.G. Cochran, *Sampling Techniques*, John Wiley and Sons, New York, 1977.
- [10] C. Constantinescu, "Statistical Techniques for Estimating Fault Coverage Probabilities," Technical Report, Duke University, 1994.
- [11] C. Constantinescu, "Estimation of the Coverage Probabilities by 3-Stage Sampling," *Proc. Annual Reliability Symposium*, Jan. 1995, pp. 132-136.
- [12] C. Constantinescu, "Using Multi-Stage and Stratified Sampling for Inferring Fault Coverage Probabilities," *IEEE Trans. Reliability*, vol. 44, no. 4, Dec. 1995, pp. 632-639.
- [13] M. Cukier, "Estimation of the Coverage of Fault Tolerant Systems," Doctoral Dissertation, National Polytechnic Institute of Toulouse, France, LAAS-Report no. 96290, Jul. 1996.
- [14] M. Cukier, J. Arlat and D. Powell, "Frequentist and Bayesian Estimations with Stratified Fault Injection," *Proc. DCCA-6*, Mar. 1997, pp. 38-57.
- [15] W. Daehn, "Fault Injection Using Small Fault Samples," *Journal of Electronic Testing: Theory and Applications*, vol. 2, 1991, pp. 191-203.
- [16] J.B. Dugan and K.S. Trivedi, "Coverage Modeling for Dependability Analysis of Fault-Tolerant Systems," *IEEE Trans. Computers*, vol. 38, no. 6, June, 1989, pp. 775-787.
- [17] J.M. Hammersley and D.C. Handscomb, *Monte Carlo Methods*, Methuen and Company Ltd., London, 1964.
- [18] P. Jalote, *Fault Tolerance in Distributed Systems*, PTR Prentice-Hall, Englewood Cliffs, New Jersey, 1994.
- [19] B.W. Johnson, *Design and Analysis of Fault Tolerant Digital Systems*, Addison Wesley, New York, 1989.
- [20] N.L. Johnson and S. Kotz, *Distributions in Statistics - Discrete Distributions*, John Wiley and Sons, New York, 1969.
- [21] N.L. Johnson and S. Kotz, *Distributions in Statistics - Continuous Univariate Distributions -1*, John Wiley and Sons, New York, 1970.
- [22] J.C. Laprie, "Dependable Computing and Fault Tolerance: Concepts and Terminology," *Proc. FTCS-15*, pp. 2-11 June 1985.
- [23] J.C. Laprie, *Dependability: Basic Concepts and Terminology - In English, French, German and Japanese*, Springer-Verlag, Vienna, 1992.
- [24] A.M. Law and W. D. Kelton, *Simulation Modeling and Analysis*, McGraw-Hill, Inc., New York, 1991.
- [25] S.V. Makam and A. Avizienis, "ARIES 81: A Reliability and Lifecycle Evaluation Tool for Fault Tolerant Systems," *Proc. FTCS-12*, 1982.
- [26] J. McGough, "Effects of Near-Coincident Faults in Multiprocessor Systems," *Proc. 5th IEEE/AIAA Digital Avionics Systems Conference*, Nov. 1983, pp. 16.6.1-16.6.7.
- [27] K.W. Miller, L.J. Morell, R.E. Noonan, S.K. Park, D.M. Nicol, B.W. Murrill and J.M. Voas, "Estimating the Probability of Failure When Testing Reveals No Failures," *IEEE Trans. Software Engineering*, vol. 18, no. 1, Jan. 1992, pp. 33-43.
- [28] V.P. Nelson and B.D. Carroll, "Fault-Tolerant Computing (A Tutorial)," presented at the *AIAA Fault Tolerant Computing Workshop*, Nov. 1982.
- [29] A. Papoulis, *Random Variables and Stochastic Processes*, McGraw-Hill, Inc., New York, 1991.
- [30] D. Powell, M. Cukier and J. Arlat, "On Stratified Sampling for High Coverage Estimations," *Proc. EDDC-2*, Oct. 1996, pp. 37-54.
- [31] D. Powell, David, E. Martins, J. Arlat and Y. Crouzet, "Estimators for Fault Tolerance Coverage Evaluation," *Proc. FTCS-23*, June 1993, pp. 229-237.
- [32] D. Powell, E. Martins, J. Arlat and Y. Crouzet, "Estimators for Fault Tolerance Coverage Evaluation," *IEEE Trans. Computers*, vol. 44, no. 2, Feb. 1995, pp. 261-274.
- [33] Z. Segall, D. Vrsalovic, D. Siewiorek, D. Yaskin, J. Kownacki, J. Barton, D. Rancey, A. Robinson and T. Lin, "FIAT - Fault Injection based Automatic Testing Environment," *Proc. FTCS-18*, 1988, pp. 102-107.
- [34] D.T. Smith, B.W. Johnson, J.A. Profeta and D.G. Bozzolo, "A Method to Determine Equivalent Fault Classes for Permanent and Transient Faults," *Proc. RAMS*, Jan. 1995, pp. 418-424.
- [35] D.T. Smith, B.W. Johnson, N. Andrianos, and J.A. Profeta III, "A Variance Reduction Technique via Fault Expansion for Fault Coverage Estimation," *IEEE Transactions on Reliability*, vol 46, no. 3, Sep. 1997, pp. 366-374.

- [36] J.J. Stiffler and L.A. Bryant, *CARE III Phase III Report - Mathematical Description*, Contr. Rep. 3566, NASA, Nov. 1982.
- [37] A. Stuart and J. K. Ord, *Distribution Theory*, vol. 1, Edward Arnold, London, 1987.
- [38] K.S. Trivedi, *Probability and Statistics with Reliability, Queueing and Computer Science Applications*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1982.
- [39] K. Trivedi, J.B Dugan, R. Geist and M. Smotherman, "Modeling Imperfect Coverage in Fault-Tolerant Systems," *Proc. FTCS-14*, June 1984, pp. 77-82.
- [40] H.M. Wadsworth (editor), *Handbook of Statistical Methods for Engineers and Scientists*, McGraw-Hill, New York, 1990.
- [41] W. Wang, K.S. Trivedi, B.V. Shah, and J.A. Profeta III, "The Impact of Fault Expansion on the Internal Estimate for Fault Detection Coverage," *FTCS-24*, June 1994, pp.330-337.

Comparing the Safety Criteria of IEC 61508 and UL 1998

J. Janeri H. Cox B. Godwin
Underwriters Laboratories Inc.
Programmable Systems Certification Services
12 Laboratory Drive, P.O. Box 13995
Research Triangle Park, NC 27709-3995

ABSTRACT

There are a vast number of products and components being developed today that rely on microprocessors and software to deliver functionality and features as well as safety control measures. These products and components need to be designed, constructed and tested according to a well-defined and practical set of safety requirements. Specifically with regard to the safety-related software, UL published a newly revised Second Edition of UL 1998 titled "Standard for Safety for Software in Programmable Components," and the International Electrotechnical Commission has developed IEC 61508, "Functional safety: safety-related systems." We describe some of the similarities and differences between these two safety standards revealing where UL 1998 is particularly well-suited for embedded safety-related software.

Introduction

Manufacturers are using more and more innovative technologies to create components, products and systems with enhanced capabilities and improved performance. When these technologies involve programmable or computerized components, benefits such as reduced parts costs and the ability to rapidly incorporate new features have been realized. With the increased reliance on microprocessors and software in safety-related systems, it is imperative that the computerized components and subsystems be subject to a standard set of safety requirements that are appropriate for this kind of technology. This, of course, immediately gives rise to several interesting questions concerning similarities and differences

between various standards and the assessment practices. In this paper, we examine two particular safety standards, namely UL's "Standard for Safety for Software in Programmable Components" [UL1998] and the international standard entitled "Functional safety of electrical, electronic programmable electronic safety-related systems" [IEC 61508], from a number of different perspectives.

From a standards perspective we note some basic similarities such as a shared set of safety objectives, similar approaches to formulating consensus, comparable language and composition/integration, and many of the same technical references sources. They are compatible in the sense that each standard has a useful scope of application – albeit at different ends of the system spectrum – and both standards consolidate software with certain microelectronic hardware requirements.

Organization

This paper is organized into two main parts: first, we will briefly explain the structure and layout scheme for both the international standards and the UL standards, providing insights into how they are developed and applied. In particular, we explain how programmable electronic system safety issues can be addressed both at the component "building-block" level and from the perspective of decomposing safety-related systems into constituent subsystems and components. Next, we highlight a few key software requirements in each standard, exposing some of UL 1998's most distinct and unique requirements for embedded safety-related software.

Comparing the Structure

The manner in which UL and IEC standards are organized, developed, and presented is indeed different. Yet after more careful study, it is clear that both strategies can be (and are) used to achieve safety assessment at the "system level." To understand the motivation behind IEC 61508 and how it works with the other IEC standards, we will first take a quick look at the method used by the IEC to organize and present its standards. In contrast, a similar explanation of UL's standards and how UL 1998 fits in to the overall standards medley will follow this discussion.

IEC Standards Structure

Figure 1 illustrates how the IEC Standards are organized into four basic facets: 1) General Requirements, 2) Collateral Standards, 3) Particular Standards, and 4) Performance Standards.

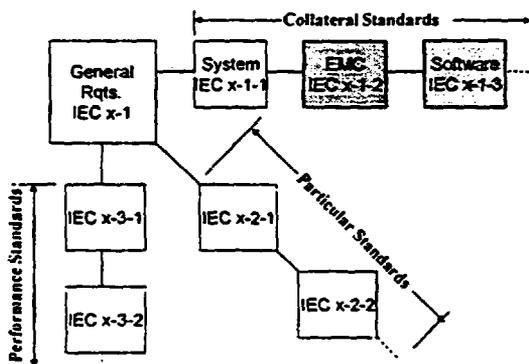


Figure 1: IEC Organization of Standards

General requirements are a collection of all of the currently known requirements that broadly apply to devices of a particular industry. So, for example, IEC 601-1 contains the core set of general requirements ("dash one") for medical electrical equipment (where "601" designates the medical device industry). Collateral Standards partition the various aspects of the system based on the technology being used in the system. These include the System, Software, Hardware, EMC, etc. The requirements in a Collateral Standard are also

general in nature and are separated out¹ so that specialists having technology-specific expertise can contribute their knowledge by way of participation on relevant IEC technical committees. Particular Standards contain device-specific requirements that focus on a given kind of device. Some examples include, light curtains, PLCs, and two-hand controls. Particular Standards hold precedence over the Collateral Standards as they may contain requirements that amend, delete, or supercede the general requirements. Performance Standards are reserved for functional and performance requirements including specific accuracy, efficiency and precision requirements. To date, very few Performance Standards have been written; instead, technical committees have folded these requirements into the Particular Standards.

The way the standards are organized allows both standards writing bodies to maintain a clean separation between the appropriate domain-specific criteria for particular kinds of technologies and devices used in a given system and the system itself as a whole.

Dual-Use of IEC 61508

How then does IEC 61508 fit into the IEC scheme for organizing standards? One way that IEC 61508 can be used as a "stand-alone" standard. This is because the requirements contained in the standard can be used directly as generic requirements (i.e., without attempts to guide the development of more refined and application specific sector standards). This role bears a strong and comparable relationship to UL 1998's role as a "Reference Standard." (Reference Standards are described in more detail in the section entitled "Structure of UL Standards").

The second way IEC 61508 can be used is as a basis for creating or generating other sector specific standards (Figure 2). The first four parts of IEC 61508 have a special status in the IEC standards classification system, known as IEC Basic Safety Publications. Per IEC policy, this means "whenever applicable [it is the responsibility of the technical

¹ The plan is that — at some point in time in the future — the Collateral Standards will be absorbed into the General Requirements.

committee] to make use of Basic Safety Publications in the preparation of its own publications.”

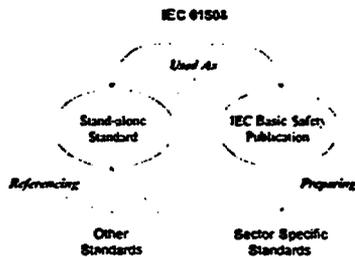


Figure 2: Dual-Use of IEC 61508

One interesting exemption to IEC 61508’s Basic Safety Publication status is the absolution of so-called “low-complexity systems.” This term is officially defined as an “E/E/PE [(Electrical/Electronic/Programmable Electronic)] safety-related system (see 3.2.6 and 3.4.1) in which:

- (a) The failure modes of each individual component are well-defined; and
- (b) The behaviour of the system under fault conditions can be completely determined.”

If taken literally – and depending on one’s definition of “system” and “completely determined” – only the most trivial E/E/PE safety-related systems would meet these criteria. Here, we use the IEC 61508 definition of *system* (Part 4, 3.3.1) which has a broad scope, meaning an individual component, a collection of interrelated/integrated components, subsystems, systems as well as a system-of-systems.

Fortunately, the IEC has stated that this base definition of low complexity system is merely a starting point and that individual technical committees may choose to select a different group of criteria that would place a more practical balance on the definition. It is clear that adjustments to the definition of low-complexity system in this

direction would increase the number and type of system that would be considered “low-complexity” and thus tend to reduce the scope of applicability of IEC 61508.

Although IEC 61508 is not yet officially published, it is being used.³ At the time of this writing, IEC 61508 Parts 2, 6, and 7 are Committee Drafts for Vote (CDV) and Parts 1, 3, 4, and 5 are Final Draft International Standards (FDIS).

We will return to IEC 61508 to examine more closely the safety-related software requirements in Part 3. For now at least, it should be clear that IEC 61508 is a general set of *system safety requirements* that are intended to be applied to E/E/PE safety-related systems and subsystems (i.e., including E/E/PE safety-related subsystems that comprise the E/E/PE safety-related system). Before IEC 61508 can be used, the application of the system must be defined. We note that since system level requirements form an intrinsic part of the standard, IEC 61508 can be applied to very large and complex systems.

UL Standards Structure

For many years, UL has developed safety standards for a wide variety of products, devices, equipment, and components. There are really only two main types of UL Standards, namely, End-Product Standards and Reference Standards (see Figure 3).

UL Standards typically only state essential requirements and do not offer guidance.

The End-Product Standards are typically very specific to a particular kind of device. For instance, UL 372 is the standard for “Primary Controls in Gas and Oil-Fired Appliances.” However, in some cases, UL standards are written for a broader scope of application. These standards, which are also considered End-Product Standards, tend to address “categories” of equipment like UL 508, “Industrial

³ Interestingly to note that the IEC only “encourages” the use of the latest published edition of an IEC Standard. In contrast, UL requires that the latest published edition with updated revisions be used.

Control Equipment,” and UL 1950, “Information Technology Equipment.”

account of views and opinions from the general public and “all interested parties.” (For more information about UL’s Standards, see [Bushell]).

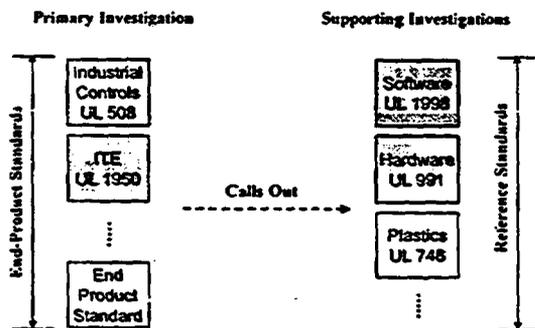


Figure 3: UL Organization of Standards

UL’s End-Product Standards are similar to the IEC’s Particular Standards.

When requirements can be collected together in documents to further support certain technology areas pertinent to many “end-products” and the assessment of products in “categories,” Reference Standards are written. The two examples that we will discuss in this paper include UL 1998, “Software in Programmable Components” and UL 991, “Tests for Safety-Related Controls Employing Solid State Devices.” Another important UL Reference Standard is UL 746 (Plastics). Although Reference Standards can be made to be self-contained, UL 1998 is not considered a “stand-alone” standard.

UL’s Reference Standards are similar to the IEC’s Collateral Standards.

Similar to the IEC technical committees, UL standards are developed using an open standards development process that follows the American National Standards Institute (ANSI) rules. UL standards are written with input from a wide range of experts; a large part of the representation comes from manufacturers who produce products that will be affected by the standard. Thus, standards organizations develop consensus standards that take

Singular Use of UL 1998

Like many other generic standards, which have broad-based applicability, UL 1998 can be tailored for a particular product or category. For instance, the typical UL framework of standards for a programmable electronic system would (minimally) incorporate the following:

1. End-Product Standard
2. UL 1998 (Safety-related software)
3. UL 991 (Safety-Related Hardware Reliability)

UL 991 is referenced by the End-Product Standard to establish estimated reliability of hardware components and to evaluate and test whether critical components meet the End-Product requirements. A failure modes and effects analysis (FMEA) is conducted on microelectronic hardware to determine the failure modes and how hardware faults can impact safety.

This is a simplified example to illustrate how the basic product safety requirements are coupled with the Reference Standards. The overall assessment will address software, microelectronic hardware requirements in addition to the product requirements. This compositional strategy also serves as a natural basis for dividing-up the supporting investigations within the Laboratories.

To summarize the UL approach, it starts with “end-products” or “categories” of products and contains detailed device-application specific requirements that have been deemed most practical to reach an acceptable level of safety. These product safety standards are akin to some combination of IEC Particular Standards (and in some cases, General Requirements and Performance Standards). The Reference Standards are most like the Collateral Standards but as with many UL Standards, they only contain requirements in mandatory language. With respect to the IEC scheme, UL standards organization can be considered (technically

speaking) on-par with and sometimes complimentary to IEC standards.

IEC 61508 Organization

IEC 61508 has seven “parts,” the first four contain generic requirements with respect to electrical/electronic and programmable electronic safety-related systems (E/E/PE SRSs). Table 1 lists the Parts by name along with the number of pages to give a sense of the relative size of each part.

| Part | Subject | Pages |
|------|-------------------------------------|-------|
| 1 | General requirements | 58 |
| 2 | E/E/PE safety-related systems | 55 |
| 3 | Software | 44 |
| 4 | Definitions and Abbreviations | 25 |
| 5 | SIL Determination Methods | 26 |
| 6 | Guidance in applying Parts 2 and 3 | 72 |
| 7 | Overview of techniques and measures | 107 |

Table 1: Parts of IEC 61508

The first four parts of IEC 61508 are normative, while the remaining parts are informative. This is not to say that Parts 5, 6 and 7 are not important. In fact, these parts of the standard give users extremely valuable guidance on developing Safety Integrity Levels (SILs), for recommending measures for E/E/PE SRSs (i.e., hardware and software measures), and for providing descriptions of various methods and techniques.

If requirements are viewed as a list of ingredients, then the informative parts (guidelines) of IEC 61508 are like a recipe that one may follow to show compliance.

UL 1998 Organization

UL 1998, first published in 1994, is now in its Second Edition and has been in use since 1995. It is a compact set of normative requirements with fifteen sections (or clauses). Appendix A contains a table of example measures to address microelectronic hardware failure modes along with a worked example and a list of brief descriptions of the measures. The appendix is normative if called

out as such by the End-Product Standard. Like other software standards (including IEC 61508) it features clauses covering definitions, documentation, configuration management, documentation control and general process requirements. It places an emphasis on risk analysis and documented traceability evidence to verify that the identified risks have been addressed. UL 1998 carves out a distinctive -- yet practical -- set of assessable testing requirements.

| Clause | Heading/Subject |
|--------|--|
| - | Forward |
| - | Preface |
| 1 | Scope |
| 2 | Definition of Terms |
| 3 | Risk Analysis |
| 4 | Process Definition |
| 5 | Tool Validation |
| 6 | Software Design |
| 7 | Critical and Supervisory Sections of Software |
| 8 | Measures to Address Microelectronic Hardware Failure Modes |
| 9 | Product Interface |
| 10 | User Interface Design |
| 11 | Software Analysis and Testing |
| 12 | Documentation |
| 13 | Off-The-Shelf Software |
| 14 | Software Changes and Document Control |
| 15 | Identification |
| - | Appendix A |

Table 2: Inside UL 1998

As shown in Table 2, the Reference Standard has been engineered to enable it to:

1. link-up with end-products (e.g. “Product Interface”);
2. interface with human operators and maintenance personnel (e.g. “User Interface”);
3. address failure modes of the underlying microelectronic hardware (e.g., “Measures to Address Microelectronic Hardware Failure Modes”);
4. address vendor-supplied software (e.g., “Off-The-Shelf Software,” and “Tool Qualification”).

Basis for Determining Measures

The use of an "index" to select among a group of safety measures is well established in the standards community. First, the index is qualitatively or quantitatively "calculated" in terms of risk/criticality and then it is used to resolve between two or more substantially equivalent safety measures. The end result is a determination of the type and the extent of the safety measure(s) to be applied.

Philosophically both standards leverage off of a similar risk-based approach for arriving at an index. IEC 61508 explicitly uses safety integrity levels (SILs), while UL 1998 uses a hazard-based approach for determining acceptable measures.

SILs require that risk analysis and reliability studies be conducted at the system level and with a known application. Without a defined application, the risks associated with some products and components can not be fully ascertained. Take for instance a programmable logic controller (PLC). It is a component that IEC 61508 regards as unevaluable until it is integrated into a system with a specific application domain to give it context. This fact restricts IEC 61508 to systems with defined applications are known *a priori*.

UL 1998 uses a simpler notion of risk analysis and "software class" to determine the acceptable measures. The difference is that UL 1998 explicitly calls for a risk analysis ([UL 1998, Clause 3]) as part of the requirements and marries together the results of the risk analysis with the notion of software class in order to arrive at suitable protective measure

Application of Standard Safety Criteria

IEC 61508 is applied to an E/E/PE safety-related system as well as any of the subordinate systems that comprise the overall system. This is done in the context of a defined application and by repeated decoupling and decomposing the system into its constituent parts, evaluating each subsystem in turn.

In some ways, it is easiest to think of UL's approach as being the opposite. Manufacturers come to UL for certification of their products and components, and each submittal is individually assessed against the relevant applicable standards. If the product or component meets these requirements, it gets one of UL's marks. The UL system and the use of independent third-party testing laboratories has traditionally been used by manufacturers that mass-produce products.

But one manufacturer's product is another manufacturer's (or system integrator's) component. So not only is "product" a relative term, but so are "system" and "component" It really depends on where you stand as to whether you can consider these as distinct and unique terms or whether they are just synonyms.

Bottom-Up and Top Down

Though there are differences in these two standards, both documents can mutually serve the manufacturing community in providing a total system safety assessment. With UL 1998, manufacturers can gain confidence in engineering systems from certified programmable systems, products and components. Of course, software is only one aspect of the overall UL assessment process. In conjunction with the UL system, manufacturers can use IEC 61508 once the system is integrated and has a defined application. We are fortunate to have this type of holistic synergism emerge in the combined US and International system safety standards.

Scope of Applicability

In the context of standards, there are at least two aspects of *scope* that can be discussed: the first is the scope of the requirements, i.e., the breadth and depth of what the requirements address. This element is covered in the next section dealing specifically with the software content of the standards. The second aspect of scope deals more with the *applicability of the standard*. UL 1998 addresses embedded non-networked safety-related software and contains a minimal number of

requirements that address microelectronic hardware. The view taken by UL 1998 is that software is merely a component of the system. That is, the requirements in UL 1998 are specifically focused on software that is embedded in the particular product. System considerations, such as how safety measures are implemented in a composite system that contains multiple subsystems, are dealt with in UL's end-product or category standards.

UL 1998 focuses solely on embedded safety-related software.

IEC 61508 is a system standard that is used by other standards (as a reference) or as a basis for developing an industry-specific (sector) standard. It contains system, hardware and software requirements in its Parts. By the scope statements in IEC 61508 and with the current definition of LCS, it can be applied to potentially any component, product, subsystem, system or system of systems.

In principle, UL 1998 and IEC 61508 treat safety-related software the same, namely they both include software implementing safety functions as well as any software that is used to develop the safety-related system (see [61508-3, 1.1.b]). However, in practical terms, the two standards differ slightly. IEC 61508 renders all of its requirements on supporting (e.g., vendor-supplied) software, whereas UL 1998 contains special subclauses (see [UL1998, 5 and 13]) to address off-the-shelf software and tool qualification.

Comparing the Software Requirements

As previously mentioned, it is difficult to compare – in a completely fair manner – UL 1998 software requirements to those in IEC 61508-3 alone⁴. This is because Part 3 is used in conjunction with Parts 1 and 2 to mold together a consistent system safety

⁴ For example, given a general requirement from Part 1, readers will find examples of additional requirements (or elaborations of requirements) addressing the same basic safety assurance theme in Part 3. A Part 3 elaboration is simply a requirement that was originally articulated in Part 1, but is now restated in Part 3 to address a software nuance.

approach to requirements writing. We believe, however, that given the coarse level of granularity with which we attempt this comparison, that there is little chance of misrepresenting any similarities or differences that may or may not exist between these two sets of software requirements.

Both standards contain requirements that address various aspects of safety-related software including: design, process, configuration management, change control, testing, and documentation requirements. We highlight a few of the more subtle differences that we found in our study, but we are cautious in drawing any hard conclusions about which set of requirements is most effective in developing safety-related software.

Software Life Cycle

IEC 61508 defines a software safety life cycle as an integral part of the standard. Clause 7 and the various subclauses form over fifty percent of Part 3 (twenty-four pages of the forty-four total). It closely follows the classic “V-model” with design, development and implementation phases being mirrored by verification and validation phases.

UL 1998 approaches software development life cycle requirements specification as set of eight high-level criteria [1998, Subclauses 4.1-4.8]. In more subtle ways, UL 1998 presents its topics grouped in such a way as to suggest a progression from the earliest phases of a software life cycle to the latest stages. Refer to Table 2, which lists the contents of UL 1998. The intent is to have a well-defined, documented and repeatable process in place. The particular clauses in UL 1998 [UL, 4.1-4.8] are derived from several sources including IEEE 1224-1994, ISO/IEC 12207, and FDA Reviewer Guidelines.

Partitioning

One way that UL 1998 emphasizes the need to reduce complexity in embedded safety-related software is by enforcing partitioning requirements. Software partitioning is the separating of software-related functions that address distinct concerns and have distinct roles. Partitioning is crucial in

embedded systems since the safety-related code is often co-resident with non-safety-related code. In addition, during execution of the computer program, memory locations, buffer caches and registers are likely to store or temporarily hold critical data that must not be corrupted. The partitioning of safety-related software from all other non-safety-related software is a principal concern in UL 1998. IEC 61508 has a somewhat different approach regarding partitioning [IEC, 7.4.2.7-8], as there are no specific testing or analysis requirements to see that the partition is enforced. The relaxed IEC 61508 partitioning requirements makes sense because it is not quite as narrowly scoped as UL 1998.

Configuration Management

IEC 61508 has seven subclauses ([IEC, 6.2.2, 6.2.3 (a-f)]) that address software configuration management. It is curious that these clauses reside in the normative portion of Part 3, yet they are prefaced by the statement: "Software configuration management should: ..." (Given that "should" is a recognized term to indicate guidance, it is not clear whether the CM requirements are mandatory or not). IEC 61508 does state that CM controls be in effect "throughout the software safety lifecycle."

UL 1998 addresses CM and change management controls in three clauses of the standard: "Configuration Management Plan" [UL, 12.4.1, 12.4.2 (a-c)], "Software Changes and Document Control" [UL, 14.1-14.5] and "Identification," [UL, 15.1-15.4]. It does not specify the duration of CM controls.

Although we have only covered a subset of the areas that are addressed in UL 1998 and IEC 6508, it should give the reader a better view of how the safety-related software requirements are directly compatible for a certain class of systems. UL 1998 is complimentary for components (UL 1998) and large systems-of-systems (i.e., IEC 61508)

Summary

Both sets of software safety requirements:

- Share similar objectives and many of the same technical references sources
- Have many of the same basic intents/comparable notions of safety
- Leverage off of a similar risk-based approach to choosing satisfactory safety measures
- Are developed by long-standing reputable organizations with consensus an input from multiple parties.
- Consolidate software with certain microelectronic hardware requirements
- Can be referenced by other standards; public domain documents

The top two differences that we believe exist between UL 1998 and IEC 61508 are:

- 1) the degree of specificity in requirements relating to development life cycle in IEC 61508 versus the process definition criteria stated in UL 1998.
- the use of safety integrity levels in IEC 61508 versus risk assessment-based approach of UL 1998.

Acknowledgments

The authors would like to thank: Mr. John Calvert for inviting UL to participate in this forum and for giving us the opportunity to compare and contrast IEC 61508 and UL 1998; Ms. Janet Flynt and Mr. Manoj Desai for their review of this paper.

References

[IEC61508] "Functional safety of electrical/ electronic/ programmable electronic safety-related systems, CEI - IEC 61508," International Electrotechnical Commission, Geneva, Version 4.0, 1997.

⁵

Interesting to note that the IEC only "encourages" the use of the latest published edition of an IEC Standard. In contrast, UL requires that the latest published edition with updated revisions be used.

[UL1998] "Standard for Safety for Software in Programmable Components, UL 1998, Second Edition," Underwriters Laboratories Inc., Research Triangle Park, NC, May 1998.

[IEEE] "IEEE Standard for Software Safety Plans, IEEE 1224-1994," The Institute for Electrical and Electronics Engineers, New York, 1994.

[12207] "Information technology – Software life cycle processes, ISO/IEC 12207(E)," International Organization of Standardization, Geneva, 1995.

UL 508, UL 746, UL 991, UL 1950, are Published Standards © Underwriters Laboratories Inc., Northbrook, IL.

[Bushell] Bushell, Beth and Sonya Bird, "Setting the Standards," On The Mark (A UL Quarterly Publication, Vol 3., No. 2, Underwriters Laboratories Inc., Research Triangle Park, NC, <http://www.ul.com/about/otm/otmv3n2/labdata.htm>, Summer 1997.

Putting Principles into Practise: The Formal Development of a Theorem Prover

Terje Sivertsen

OECD Halden Reactor Project

P.O. Box 173, N-1751 Halden, Norway

Abstract

Formal methods for software specification, verification, and development have for many years been a focal point of research at the OECD Halden Reactor Project. A major accomplishment has been the establishment of a complete methodology for the practical application of algebraic specification in formal software development. The methodology is supported by the HRP Prover, an automatic theorem prover developed at the Halden Project to facilitate exploration of animation and theorem proving techniques in formal software development. The paper presents results from a project involving the use of this methodology in the development of a new version of the HRP Prover. In spite of the large number of tools available that support formal methods, very few of these have been developed in accordance to the same principles. The project has delivered a tool that facilitates formal development of modular, well-structured programs by the use of automatic program generation from specification or design. As a consequence, functional requirements can be realized directly, while improving the traceability of requirements from the program code. This also simplifies maintenance and further development, since new program code can be constructed directly from changed or added requirements. The reported results from the development project are of relevance to the formal development of a wide range of language-oriented tools, involving aspects like analysis, transformation, and code generation. In particular, the approach employed appears to facilitate combination of complementary specification notations. This is exemplified in the paper by the integration of Petri nets and algebraic specifications.

1 Introduction

The recent years have witnessed the replacement of many conventional electro-mechanical process control systems with computer-based systems. This also includes the use of computers in safety-related tasks, e.g. in nuclear power plants and traffic control. The motivation behind this shift towards the use of programmable equipment is manifold. Important benefits are the possibilities for implementing more accurate trip criteria, the improved means for automatic surveillance, as well as simplification of calibration and functional testing during operation. There are however also more pragmatic concerns relating to decreasing availability of equipment and spare parts for the conventional systems and of personnel with appropriate technological expertise. Nevertheless, there has been a certain reluctance to the use of programmable equipment in

safety systems. One reason for this reluctance has been the complexity of safety assessment and licensing of these systems, in particular of the embedded software.

Recently, the application of formal software development methods have been treated with increasing interest within the nuclear society. Much discussion and, to a certain degree, controversy arose from the verification and validation of the computer-based Darlington shutdown system [4]. Nevertheless, there is today a growing consensus within the nuclear society that more practice on the use of formal methods is needed in order to evaluate their applicability [25]. Several independent studies suggest that there is a need for a systematic, rigorous effort in establishing design requirements to minimize errors in the final product [1]. Licensing authorities in general have a particular interest in representative applications of existing formal methods to make decisions on whether the use of formal methods should be required, which formal methods should be used, what is the appropriate way to use them, and what to require to be formally verified. Much of this motivation comes from the limited value of traditional methods. Following [14], "traditional software-development techniques usually do not provide the levels of dependability demanded by safety-critical systems, and the quality criteria are usually such that the amount of testing that is feasible cannot demonstrate that the desired goals have been achieved". As a matter of fact, there are several important aspects that make the application of software in safety-critical applications fundamentally different from their application in other areas. Safety-critical applications must work when needed, and it is not appropriate to wait for evaluation during use to bring the reliability up to an acceptable level. The realization of the potential benefits of computer-based control and safety systems for nuclear power plants therefore requires *verifying* the reliability of these systems. Traditionally this has been done by means of simulation of the hardware design and exhaustive software testing. It appears however that the use of formal mathematics, in some form, is necessary in order to achieve substantial improvements in the development of dependable software.

The present paper focuses on the formal development of tools supporting formal methods. There is today a wide variety of tools available that provides such support. Nevertheless, it is regrettable that very few research and development activities have given serious attention to the formal development of these tools. It is believed that the eventual maturity of formal methods will require a change in this attitude. That is, formal methods should not only be used in the development of conventional software systems, they should also be used in the development of formal methods support tools. This also has a practical aspect, since the development of systems requiring a high level of safety would require development tools with very high reliability. According to [6], the combined set of tools used in the development of safety system software shall provide the same level of dependability as the level required from the target software.

Formal methods for software specification, verification, and development have for many years been a focal point of research at the OECD Halden Reactor Project. A major accomplishment has been the establishment of a complete methodology for the practical application of algebraic specification in formal software development. The methodology is supported by the HRP Prover, an automatic theorem prover developed at the Halden Project to facilitate exploration of animation and theorem proving techniques in formal software development. The HRP Prover supports a methodology that allows for using the same language, tool and proof techniques both in specification and design, even down to a "concrete" specification. In the specification phase, the HRP Prover is used to verify and validate the specification, while in the design phase the same tool is used to verify the correctness of the design steps. A few years ago, the applicability of the methodology was demonstrated in a case study on the development of a reactor safety system [21]. Since 1995, the methodology has been used in the development of a new version of the tool. That is, the tool is developed in accordance to the same principles as it is intended to support. The existence of the tool provides evidence to the claim that the methodology based on algebraic specifications and the HRP Prover can be used effi-

ciently in the development of programs of realistic size and complexity. To the knowledge of the author, the new HRP Prover is the first tool of its kind that has been formally developed in accordance to the same methodology as is supported by the tool. It appears to be unmatched from earlier developments of theorem provers in its extensive use of algebraic specification.

The paper is organized as follows. Section 2 presents findings from the EvalFM project, where algebraic specifications and the HRP Prover were used in a case study on the formal development of a reactor safety system. Section 3 introduces the formal development of the HRP Prover, and presents some of the main principles employed. Section 4 discusses central aspects of the specification of the HRP Prover, including the specification language, the concept of abstract syntax, the various levels of theorem proving, and finally transformation and code generation. Aspects related to implementation and maintenance of the generated program are discussed in section 5. A useful new feature of the HRP Prover is the automatic transformation between *state-based* and *transition-based* specifications. Section 6 describes the two classes of specifications, their application in the combination of Petri nets and algebraic specifications, and finally their relative merits in software specification and design. The issue of software quality is emphasized in section 7, where the implications of the development process on various product quality aspects are discussed. Finally, section 8 considers the transferability and general relevance of the approach employed in the formal development of the HRP Prover.

2 A Case Study on Formal Development

The objective of the EvalFM Project [21] was to evaluate the applicability of formal methods, and in particular the HRP Prover, in the development of a realistic, preferably a real, safety-critical system related to nuclear power plant operation. In co-operation with Sydkraft and ABB Atom in Sweden, a case example was defined on basis of the computer-based power range monitoring (PRM) system installed at Barsebäck NPP. The case study did not address ABB's implementation of the example system, but the development of a similar system using formal methods. The formal specification was based on the original customer's requirements document for the system, and was independent of ABB's implementation. The purpose of the PRM system that was of particular interest in this case study was the monitoring of the average power emission of the core. When high power emission is monitored, the system must trip the high level alarms. Based on the requirements document, the EvalFM project produced a formal algebraic specification of one out of four similar subsystems of the PRM system, utilizing a general mathematical tool-kit defined for the method. Finally, the subsystem was designed and implemented in a subset of Pascal. The case study also investigated how the design could be varied to put stronger emphasis on efficiency. The results provides clear evidence to the claim that formal methods can be utilized in the development of a real safety-critical system. At the same time, it was concluded that the potentials of formal methods would increase whenever the customer's requirements document allows a higher flexibility with respect to design and implementation. The development method based on algebraic specification supported in a natural way the implementation of a program that avoids potentially dangerous features of the Pascal language.

An important aspect of the specification process was the derivation of the abstract functionality from the technical descriptions provided by the customer. This would however be necessary whether or not a formal specification language was chosen, as the requirements document describes the desired system in a way which apparently suggests a specific, analogue hardware implementation. There were however important non-functional requirements for which the usefulness of algebraic specification, as well as of formal specification languages in general, is very limited. In the case study, this first of all related to the given requirements to technical performance and accuracies.

The use of the HRP Prover formed an important part of the development of the case example. The tool supported the detection of syntactic errors in the specification, the execution of the specification as an early prototype of the system, proofs of properties of the specification, and proofs of the correctness of the design steps. All of these activities involve a large amount of symbolic manipulation, and the provision of a powerful theorem prover is therefore essential for the success of the method. Nevertheless, the isolated use of a theorem prover would probably not be sufficient in an industrial development project. Industrial use would presumably require a smooth integration of the tool in an application-oriented environment which included the theorem prover, relevant text editors, graphical user interfaces, transformation tools, etc. This is in agreement with [6], which recommends that tools are incorporated into an integrated project support environment to ensure proper control and consistency. It would also be essential to ensure that this environment was sufficiently reliable.

The following findings in the EvalFM project were also given in [21]. Since the purpose of the project was to evaluate established techniques, the reader should find many of them familiar. The findings are noteworthy first of all because of the particular character and importance of the application domain, i.e. the formal development of reactor safety systems.

- Formal specification can be facilitated by the use of some library of pre-defined data type specifications.
- Algebraic specification can be used in the design as well as in the specification, and allows for implementations in a wide variety of programming languages.
- The potentials of formal methods are increased whenever the customer's requirements allow for a higher flexibility with respect to design and implementation.
- Formal software development supports the implementation of programs which avoid undesired features of the chosen implementation language.
- There are important non-functional requirements for which algebraic specification provides little support, such as requirements to technical performance and accuracies.
- Whenever the customer's requirements are described in terms of an analogue implementation, certain modifications are necessary in order to use the requirements as a basis for the development of a digital system. The incompleteness or incorrectness of these modifications is the source of an important class of specification errors.
- Execution of the specification is an effective means for detecting specification errors and can be performed incrementally during the production of the specification.
- Execution of the specification increases its comprehensibility, and thereby facilitates the communication between agents with widely varying technical background.
- Executable algebraic specifications appear to provide a sufficiently high abstraction level in most cases; the major limitation to the abstraction typically relates to the concrete nature of the customer's requirements.
- The assessment of an algebraic specification can be performed both by execution and by proving expected properties.
- Efficient use of theorem proving in specification and design requires that the specification language is supported by a powerful theorem prover. For safety-critical applications, parts of such a tool should probably be developed using formal methods.
- Algebraic specification supports a gradual design of the specification towards an implementation, and

provides a framework for proving the correctness of the design steps.

- Industrial use of algebraic specification, as well as of formal methods in general, would require the provision of a reliable integrated environment.
- A large part of the effort involved in a formal development project is invested in the production and assessment of the specification. In most cases, only a minor part is invested in the actual implementation.

3 The Formal Development of a Theorem Prover

The EvalFM project gave a demonstration on the applicability of formal methods and the HRP Prover in the development of safety critical software. The case study did however not give any clear indication as to the possibility of using similar principles in the development of tools supporting such methods. This possibility relates to several factors that motivated the formal development of a new version of the HRP Prover:

- Industrial use of formal methods, in particular in applications related to safety, requires highly reliable support tools. Very few support tools have to any great extent been developed using formal methods.
- Software is increasingly used in I&C systems important to safety. The anticipated variety of software in future systems requests a similar extension in the types of systems treated by formal methods.
- The formal development of the HRP Prover would be representative for the formal development of a wide variety of language-oriented tools. It was therefore expected that the research results would provide general guidance to the application of formal methods.
- A significant part of the development would focus on transformation of specifications and code generation from specifications. The research results would specifically be applicable to the development of automatic code generators for conventional programming languages.
- There is an increasing interest in applying formal methods also for systems not directly related to safety. A successful formal development of the HRP Prover would indicate that formal methods can be invested in the development of relatively large programs without incurring excessive development costs.

The novelty of the approach is seen first of all from the fact that the methodology used in the development of the tool is identical to the methodology supported by the tool. Furthermore, the HRP Prover is automatically implemented from its specification, providing a tool in coherence with its specification. Due to the non-trivial nature of the application, the development provides evidence that programs of realistic size and complexity can be developed efficiently by the use of algebraic specifications and the HRP Prover.

3.1 Principles of the Development

In the following, we will briefly describe some of the basic principles employed in the formal development of the HRP Prover. Note that the applicability of these principles is not restricted to this particular application.

Formal specification of functionality: Basic to any formal development is the provision of a formal functional specification. In the present case the specification covers functionality related to lexical and syntactical analysis of algebraic specifications, static semantics through type checking, dynamic semantics through evaluation, automated theorem proving, specification transformation and code generation. All aspects of the core functionality were specified in algebraic specification and supported by the conventional version of the HRP Prover.

Automatic code generation: Since the definition of the specification language as well as the functionality of the HRP Prover is given in the specified language, the specification of code generation could be utilized in the automatic implementation of the overall specification into an executable Prolog program. This program constitutes the major part of the new HRP Prover. Future extensions can be implemented in the same way, as mere extensions to the existing tool.

Layered, incremental development: The formal development of the HRP Prover can be described in terms of layers of functionality. By way of example, the code generation can be specified on basis of an intermediate representation of the source language, independently of functionality related to automated theorem proving. Correspondingly, automated theorem proving can be specified independently of the functionality related to code generation, but may implemented by using a limited new version of the tool featuring this functionality. In this way, the HRP Prover can be developed incrementally, where preliminary versions with less functionality may be utilized in the development of their more complex successors.

Concepts from compiler development: The principles and techniques used in the development of compilers permeate many areas of computer science and software engineering. This is also the case in the formal development of the HRP Prover. Basically, a compiler reads a program written in the *source* language, and translates this into an equivalent program in the *target* language. Of particular interest in our context is the *intermediate* language in which the source specification can undergo analysis. A clear partitioning of a compiler into a front and back end also makes it easier to make several versions of the compiler running on a variety of machines. Analogously, a similar partitioning of the HRP Prover makes it easier to define translations of algebraic specifications into different programming languages.

Abstract syntax trees: In the analysis of a program, it is common to represent the program in a tree structure, in terms of so-called *abstract syntax trees*. This corresponds closely to the hierarchical structures defined by the generator terms in algebraic specifications. Due to the inherent properties of the specification language, algebraic specification provides the constructions needed for defining and performing the analysis of the source specifications in an efficient and elegant way. The hierarchical representation of source specifications also facilitates the implementation of structure (context-sensitive) editors, pretty printers, specification transformation, code generation, etc. Especially important is the means this representation provides for defining the analysis of the terms and formulas in the source language.

Metacircular evaluation: In the definition of programming languages, it is common to define a procedure that, when applied to an expression of the language, performs the actions required to evaluate that expression. Such a procedure is called an evaluator. A popular approach is to use metacircularity, i.e. to write the evaluator in the same language that it evaluates. This approach is followed in defining evaluation of expressions in algebraic specification. This means that the design and implementation of the HRP Prover specification can be undertaken using established techniques supported by the very same tool.

4 Specification Aspects

The major part of the activities involved in the formal development of the HRP Prover relates to the provision of a formal functional specification of the tool. The appropriateness of algebraic specifications in describing the functionality can be compared to the benefits of a problem-oriented language set against those of a machine-oriented language. The functional specification consists of a hierarchical set of data type specifications and module specifications written in the algebraic specification language supported by the very same tool. In general, the provision of a formal specification improves preciseness, because it disciplines the specifier to state explicitly the information necessary to determine what is intended in a particular cir-

cumstance. Applied on the specification of the HRP Prover, this means that a sufficiently precise description is given of the syntax and semantics of the specification language, the way theorems are proven, how executable code are generated, etc. This description suffices as a basis for implementing these functional aspects of the tool, and it can also be used as a basis for ensuring that using the tool does not adversely affect the production of software. In particular, a precise specification of how algebraic specifications are translated into Prolog makes it possible to demonstrate that the use of a Prolog interpreter does not contribute to the logic of the specifications in other ways than expected.

4.1 The Specification Language

An essential part of the formal specification of the HRP Prover is a definition of the specification language supported by the tool. It needs to be emphasized that this definition is written in the very same language. This specification language has through several years been supported by the conventional version HRP Prover, and is essentially an algebraic specification language in the tradition of J.V. Guttag [8]. The specification language can in many respects be compared to a high-level, strongly typed, programming language. The language is intended for the specification of functional requirements, design, and architecture of conventional software by means of abstract data types [2]. The modularity of the language also facilitates reuse of (parts of) specifications. Reuse can be further facilitated by the establishment of libraries of generally useful specifications. This is a well-established practice in object-oriented programming, where reusability of code is facilitated by the techniques of inheritance and genericity. In a similar way, algebraic specifications can be reused by using the techniques of parameterization and enrichment:

- *Parameterization*: A data type is made generic through the use of formal parameters. Instantiation corresponds to replacing the formal type parameters with actual types.
- *Enrichment (extension)*: A new data type is introduced by extending a specification with new functions, without modifying the original specification.

The specification language is restricted to first-order functions, i.e. a function may not take functions as arguments or return a function as value. The concept of signature of a specification is adopted from universal algebra, where it means a set of function symbols with arities (i.e. number of arguments). The axioms are built as conditional equations, where all variables are first-order universally quantified over their respective types (as given by the declarations). While the *definitions* in a specification are restricted to conditional equations, the *constraints* have a more general axiomatic form.

We will conclude these remarks on the specification language by briefly touching upon some of its limitations:

- As usual within algebraic frameworks, the specification of infinite values such as *streams* is not supported. Streams refer to flows of data, usually through a channel, between a sender and a receiver.
- The specification language can not express *pollution* (introducing new values for a given type) or *collapse* (equating values which were previously distinct) in the sense of OBJ2 [7]. OBJ2 supports the generalisation of standard many-sorted algebra to so-called *order-sorted algebra*, which is motivated by the treatment of partial functions. Order-sorted algebras are not supported by the specification language of the HRP Prover.
- The notion of *subsorts* is not supported directly, and has to be modelled by the use of *embedding (or projection) functions*. These functions reflect embedding in the strict mathematical sense by converting an element of some sort into an element of another sort of which the first sort is a subsort.

4.2 Abstract Syntax

The *abstract syntax* of a specification captures the essential structure of the specification but suppresses its concrete representation. The abstract syntax is given in the form of an *abstract syntax tree*. This representation facilitates the definition of a semantics for the specification language, as well as the specification of functionality related to prototyping, transformation, etc. The abstract syntax tree differs from a *parse tree* by representing some syntactic features (parentheses, commas, periods, etc.) only implicitly by means of its structure. In the specification of the HRP Prover, the *parser* is specified with a function that takes as input the sequence of tokens output from the lexical analyser and produces an abstract syntax tree represented by an element of a type SPEC. The grammar is unambiguous in the sense of equivalence between the possible abstract syntax tree representations for any given input sentence. In our case, the grammatical phrases are terms of data types written in algebraic specification. The approach followed is so-called *predictive recursive-descent parsing*, which is based on recursion without backtracking.

Most of the specification of the parsing (also known as *syntax analysis* or *hierarchical analysis*) consists of functions that transform input sequences of tokens into abstract syntax trees. By way of example, the abstract syntax tree of a definition is specified with the data type DEFINITION. The overall specification of the parser is given by the specifications of a large, hierarchical set of data types, see Figure 1. The figure illustrates the hierarchical relationship between the data types, where e.g. the data type SEMANTIC_PART depends on the data type DEFINITION, but not the other way around. In a similar way, the data type DEFINITION depends on the data types DEFTERM and FUNCTOR, etc. For simplicity of presentation, the triangles denote subtrees of data types.

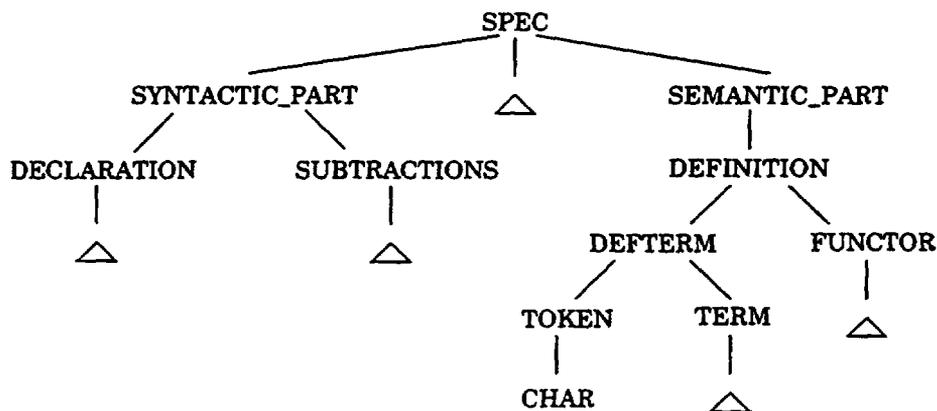


Figure 1. The hierarchical relationships between the data types.

The syntax is described by a grammar, and is therefore not sufficient for characterizing the class of acceptable specifications. An important task in *semantic analysis* is to collect type information and check the source specification for errors related to types. This process, called *type checking*, reveals errors related to the use of arguments having types different from those specified for a function. The type checking of a specification is carried out on basis of the abstract syntax tree constructed in the syntax analysis. Semantics in the full sense is in this context the meaning of a term in the specification, and is described by means of term rewriting.

4.3 Levels of Theorem Proving

The formal development of the HRP Prover involves the establishment of a uniform approach to term evaluation and theorem proving in the context of algebraic specifications. The approach involves several levels of proof, see Figure 2.

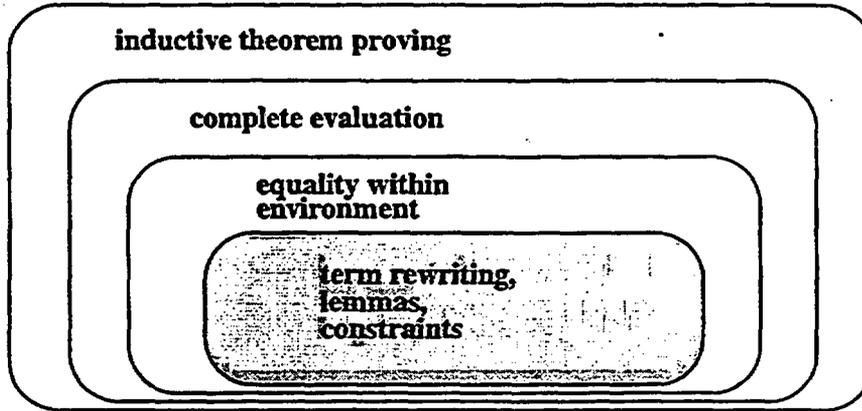


Figure 2. Four levels of theorem proving.

At the first, uppermost level, a theorem (or a set of theorems) is dynamically refined into several simpler theorems in accordance with an induction principle. For each of these simpler theorems, a complete set of cases are generated at the second level. These cases correspond to all logical combinations of assumptions dynamically asserted during the proof. Each of these combinations constitutes an environment in which the theorem is proven. This utilizes basic term rewriting, lemmas proven previously, and user-defined constraints. The approach can be used in systematically proving a large number of interdependent theorems. We will in the following look into these levels in some more detail.

The uppermost level of theorem proving is based on the author's algorithm for generating an exhaustive induction principle during the proof of a conjunction of theorems. The generation of induction principle reflects the use of *multiple induction*, which can be defined recursively as either simple induction, or multiple induction within a simple induction. Note that multiple induction may involve induction variables of different types.

Each induction variable gives rise to refinement of a theorem into several cases. In proving each of these cases, another level of cases is introduced at the second level by the dynamic assertion of assumptions made during the proof. Complete evaluation for all logical combinations of assumptions is based on a second algorithm invented by the author. This algorithm is based on making assumptions about the truth-value of undefined conditions encountered in the evaluation of a term or in the proof of a theorem. By dynamically extending and reducing a set of numbers corresponding to the rows in a truth table, a complete proof or evaluation is generated.

The refinement of theorems provides environments in which equality is evaluated. The criteria determining semantic equality of two terms is obviously an important consideration in proving equational theorems. Note that these criteria need not be complete in order to ensure sound proofs. What we are aiming for is an adequate set of criteria that can be specified in a transparent way and implemented with sufficiently efficient algorithms. At the fourth, and lowest level, terms are evaluated by means of term rewriting. Since terms are

represented by a directed graph, the technique known as *graph reduction* is applied to reduce the complexity of rewriting a term. This is done by replacing subtrees (representing subterms) by the expansion of the term they represent.

Whether or not theorem proving is supported by mechanized means, the use of lemmas remain a basic ingredient in the process. It has been demonstrated earlier how the HRP Prover supports a systematic search for lemmas providing a complete proof of the original conjecture [20]. The new HRP Prover also supports constraints, which are closely related to lemmas. Constraints and lemmas assume the same form, and are utilized by the theorem prover in the same way. An important difference related to pragmatics is that constraints represent axioms of the specifications, and are therefore not proven *per se*. Instead, the use of constraints put specific obligations on the proof of consistency of the specifications.

4.4 Transformation and Code Generation

Part of the formal specification of the HRP Prover covers automatic translation of algebraic specifications into the programming language Prolog. Since Prolog can be viewed as a descriptive as well as a prescriptive language, it suits well for many applications where formal relationships and objects play a central role. Prolog is particularly relevant for implementation of algebraic specifications because of the natural relationship that exists between definitions in algebraic specification and clauses in Prolog. The chosen implementation language is therefore highly problem-oriented in this case.

The approach chosen to translate specifications is quite similar to the compilational, innermost method of M.H. Van Emden and K. Yukawa [5], but extends this approach at several points to cope with the full specification language. The approach is called *compilational*, because terms and equations are compiled to a representation in some model. This model is the subset of first-order predicate logic representable as Horn clauses in Prolog. The approach requires that the functions of a data type are separated in the usual way into *constructors* and *defined functions*. The constructor terms are those constructed by composition of variables and generators. For each defined function, i.e. the outermost function on the left hand side of an equational definition, and the defined functions occurring on the right hand side, a corresponding Prolog predicate with the arity increased by one is introduced. Since the arguments of the left hand side are either constructor terms or variables, these will not be translated, and the arguments of the corresponding predicates can be "copied" from the specification. The remaining argument at the last position in the predicate will be a variable unified with a term representing the value of the function. An advantage with this approach is that no unification process is needed other than the one provided by Prolog. The approach makes it possible to successfully execute algebraic specifications, and to use induction to prove properties of specifications.

The specification of code generation into the programming language Prolog provides in effect an operational semantics of the specification language. This translation is generally applicable to all algebraic specifications in the defined language, and is not limited to the implementation of the HRP Prover. In fact, the translation specified at one stage of the development project was used in the new HRP Prover to implement a command providing automatic generation of stand-alone Prolog code from a given specification. This could then be utilized in the further development of the tool.

In relation to transformation tools, an important analysis task is to identify the nature of the faults these tools can introduce into the target software, together with their consequences [6]. When it comes to the transformation of algebraic specifications into Prolog, it can be demonstrated that the use of a Prolog interpreter does not contribute to the logic of the specifications in other ways than expected. Since the new HRP Prover is implemented using this transformation, such contributions could have influenced the behaviour of the tool in non-obvious ways. In fact, only a very small part of Prolog is used in implementing algebraic specifica-

tions. The Prolog interpreter is in this context mainly used as a term rewriting machine for *ground terms* in algebraic specifications. This is also reflected in the implementation of the new HRP Prover, where this is ensured by means of an interface between user input and execution of rules resulting from specifications. It is of course possible to execute these rules with open terms as well, but this is not necessary in the execution of the HRP Prover.

When applying transformation tools in the development of safety system software, accepting their output without further review would require special justification. One approach to validate the output is to use reverse engineering tools, with the aim of translating the code into a form suitable for comparison with the specifications. This would be possible with the transformation from algebraic specifications into Prolog, since the generated Prolog code in principle can be transformed back to a specification that is equivalent to the original one. This is a feature that will be considered in future extensions of the new HRP Prover. It should however be emphasised that these results are not automatically transferrable to code generation into other implementation languages.

5 Implementation and Maintenance

From a perspective of implementation size, the formal development of the HRP Prover is a relatively large application of formal methods. The basic functionality of the new HRP Prover is implemented by a Prolog program consisting of more than 1800 clauses and 14000 lines of code. More than 90% of the overall code (not counting some trivial explanatory output routines) has been formally specified in algebraic specification and automatically generated into Prolog. The remaining <10% of the code facilitates user interaction and file I/O, and was programmed directly in Prolog. The size of the program is expected to increase significantly as a result of future extensions. These extensions can be specified and implemented in a modular way using the same methodology. The issue of modular software design is important since it provides an approach to minimise the risk of design faults and to facilitate independent verifications. The use of algebraic specifications is particularly strong when it comes to recommended practices like hierarchical modularization, encapsulation and information hiding [6].

We will in this section focus on some implementation and maintenance aspects related to the approach followed in the formal development of the HRP Prover.

Layered implementation: As was indicated in section 4, the overall specification of the HRP Prover is organized in terms of a hierarchical set of data type (and module) specifications. This reflects the general strategy to the specification of the HRP Prover, viz. in terms of layers of functionality. Code generation into Prolog maintains these layers by confining its scope to a given data type or module specification, independently of other specifications or code. In spite of the apparently "flat" structure of a Prolog program, the code generated from the specification can easily be organized in terms of readily identifiable blocks of a program, maybe on separate files. In general, this means that extensions to a generated Prolog program can in every respect be implemented as mere extensions of the existing tool.

Realization of functional requirements: Automatic code generation from specifications ensures that functional requirements are realized directly, without the need for manual work in refining or implementing the specifications followed by discarding of proof obligations related to these steps. This simplifies the implementation, verification, assessment, and maintenance of a program. It is still possible to carry out design steps by refining specified data structures and operations, with the aim of providing e.g. a more efficient implementation.

Traceability of functional requirements: For many activities involving the implementation of a program, the

traceability of functional requirements is of vital importance. Verification and assessment of a program is greatly facilitated if the relationship between requirements and the various code segments is well documented. In a similar way, but with a somewhat different perspective, maintenance and possible further development of the program need to consider the same relationship in order to understand how software changes affect the correctness with respect to the specified functionality. In particular, it is essential in software maintenance that changes to the code does not incur unintentional changes to the functionality. If the code is generated directly from the functional requirements or otherwise has a well documented relationship to these, the resulting traceability of requirements facilitates a more dependable approach to software maintenance. Instead of changing code directly, the changes are made to the related requirements, which in turn are realized in new code.

In conventional practice, the specification of a software system is typically not complete before the final implementation is running. The reason is that the developers knowingly provide an implementation that redefine the specification itself [24]. Modifications to the specification also often arise after the system has been put into use. The actual usage of the system often reveals weaknesses that were not envisaged in the specification phase. It must however be stressed that in the development of safety systems, much effort is invested on achieving a correct specification in the first stage, and that relatively few modifications of the specifications are made during the development. The high safety and reliability requirements to these systems also imply that rigid restrictions are put on the possibilities to modify existing implementations. As a consequence, unexpected modifications are usually implemented in accordance to a formalized procedure.

6 State-based and Transition-based Specifications

An important feature of the new HRP Prover concerns two related classes of algebraic specifications that capture the concept of state. The *state-based* specifications model the state explicitly, while the *transition-based* specifications model the state implicitly by constructing a traditional generator basis. It can be demonstrated that specifications in each of these classes can be transformed into specifications in the other class, while preserving the proven properties of the specifications. As a practical consequence, the specifier can readily re-organize his specification so that it conforms to the most efficient or familiar approach, or to what appears to provide the best starting point for designing and implementing the specified system.

Each of the two different classes of specifications involves a particular strategy to the specification of states and transitions, see Table 1:

Table 1:

| States and transitions in algebraic specifications | | |
|--|------------------------------------|---------------------------------------|
| | <i>state-based specification</i> | <i>transition-based specification</i> |
| generators | one record of state variables | initial states, transitions (many) |
| functions | initial states, transitions (many) | state variables (many) |

In state-based specifications, the state space of a system is specified as a tuple of values, while the transitions are specified by defining functions giving new state values. In transition-based specifications, the transitions are represented by the generator symbols of a specification, and the generator terms are interpreted as "histories" of the real system. A state variable can then be specified as a function taking such a generator term

as argument and returning the value of the state variable as the system has engaged in the transitions represented by the generator term (in the given sequence). Verifying invariants for the transition-based specification is done by traditional generator induction, but that in effect gives the same inductive schemes as are used for the state-based specification. In both cases induction is done on what represents the transitions in the real system. The underlying theory clarifies when and how a notion of equivalence between these two types of specifications makes it possible to freely choose between two different approaches to specification and verification, while preserving the set of proven properties of the specifications.

In the specification of the HRP Prover, the notions of state-based and transition-based specifications are made precise by the definition of predicates checking whether the given data type specification belongs to the appropriate class. Furthermore, separate functions have been defined that transforms a given data type specification in one of these classes into the corresponding specification in the other class. These predicates and functions form the basis for the implementation of the corresponding commands in the new HRP Prover.

6.1 . Applicability for Petri Nets

The concepts of state-based and transition-based specifications have been applied in the establishment of a uniform approach to the translation of a wide variety of *autonomous* and *non-autonomous* Petri nets into algebraic specification. The importance of Petri nets in this context first of all relates to their usefulness as an intermediate language between a wide variety of graphical descriptions on one hand, and textual formal specifications on the other. Furthermore, the importance of Petri nets in the nuclear sector is well documented through applications such as fault diagnosis [11] and fault detection [17] [18] in nuclear reactors, fault tolerance in nuclear reactor protection systems [3], and modelling of work flow in nuclear waste management [15]. A characteristic of Petri nets [16] [19] is that they are at the same time state and action oriented, in the sense that both the states and the actions are explicitly described. According to [12], most system description languages describe either the states or the actions - but not both. Using Petri nets, the reader may easily change the point of focus in the course of analysing the net. A similar change of focus is feasible in algebraic specification through the transformation between state-based and transition-based specifications.

The approach involves translating Petri nets optionally into state-based or transition-based algebraic specifications, and using automatic transformation between these two classes in order to utilize their relative merits. The translation makes it possible to analyse the nets with techniques established for algebraic specification, including the use of the HRP Prover, see Figure 3.

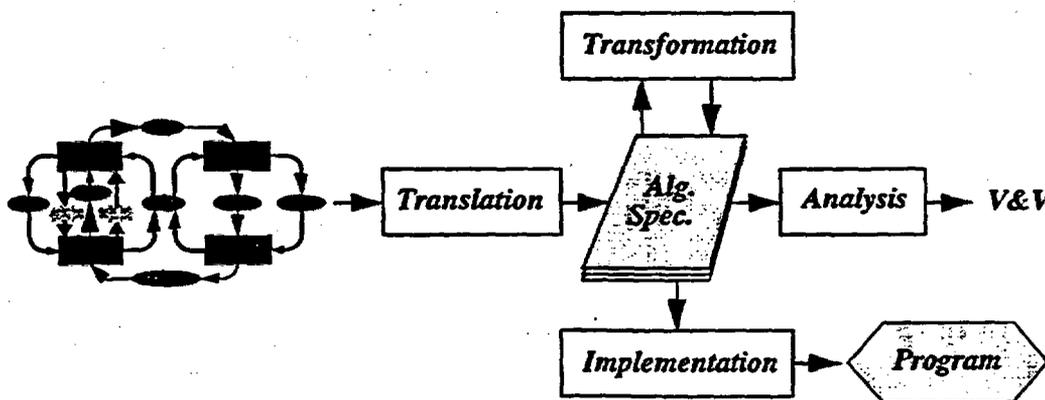


Figure 3. Translating Petri nets into algebraic specifications.

Furthermore, the nets can be manipulated indirectly through transformation of the resulting specifications. In this way, Petri nets and algebraic specifications can be integrated to form a common specification document to be used as a basis for the software development. The approach has been generalised to many different kinds of Petri nets including

- *Autonomous nets*: Condition/Event systems, Place/Transition nets, Coloured Petri nets.
- *Autonomous nets with increased descriptive power*: Inhibitor arc Petri nets, Priority Petri nets.
- *Non-autonomous nets*: Synchronized Petri nets, Timed Petri nets (weak and strong time semantics), Mixed synchronized and timed Petri nets, Interpreted Petri nets.

The choice of state-based or transition-based specification relates to pragmatic concerns about the relative merits of the two classes of algebraic specifications, both in general and with respect to the translation and extension of Petri nets. When it comes to non-autonomous nets, the introduction of synchronization favours the state-based specification, since new external events can then be included successively by introducing new functions. On the other hand, the introduction of timed places favours the transition-based specification, since timing constraints can then be added or modified successively by introducing or modifying separately defined functions. Also the introduction of data processing favours the transition-based specification, since data variables can then be added or modified by introducing or modifying separately defined functions.

Some of the most important findings from the research activities on Petri nets and algebraic specifications can be summarised as follows.

- The algebraic specification resulting from the translation can be used to prove general invariants by means of induction on the transitions.
- The algebraic specification of a Petri net can be gradually extended to incorporate new functioning rules.
- The basic approach to the translation of Petri nets into algebraic specifications can be generalised to a wide variety of Petri nets.
- Automatic transformation between state-based and transition-based specifications gives a flexible and efficient approach to the combination of Petri nets and algebraic specifications.

The work on Petri nets and algebraic specification represents a promising approach to the integration of specifications written in different languages. The approach appears to be applicable to a wide variety of graphical and textual specification languages. Furthermore, the transformation between state-based and transition-based specifications can be utilized in the combination of specifications written in different *textual* languages, such as Z notation [23] and Larch Shared Language [9]. On the basis of these observations, it is expected that the approach will prove useful in future work on combination of specification techniques.

6.2 Relative Merits of State-based and Transition-based Specifications

We have already seen that there are good reasons for insisting on the need for both state-based and transition-based algebraic specifications. Many of the differences between a state-based and a corresponding transition-based specification emerges when extending, combining, transforming or refining specifications, or when using tool support. Some of these activities may in practice demand one particular of the two specifications, and consequently a need to transform specifications to suit these needs. The need to transform specifications arise in particular when *combining* a state-based and a transition-based specifications. The style of the combined specification is determined by the desired focus in each concrete case, but the com-

ination will anyhow involve a transformation of one of the specifications.

We will in this subsection briefly summarize some of the differences between the state-based and the transition-based style of specification. The choice of specification style will for a particular case typically require a consideration of several of these. Furthermore, the relative merits represented by these differences might be weighted differently from case to case.

Information contents: An important difference between a state-based and a corresponding transition-based specification relates to the information represented in the generator terms. In a state-based specification, a generator term represents the “current value” of each state variable, but it is in general not possible to relate the current state to previous states. In comparison, a generator term in a transition-based specification provides an explicit representation of the history of the system. That is, the generator term gives the sequence of transitions in which the system has engaged, and thus represents the information needed to distinguish between the different paths leading up to a particular state.

Specification process: A state-based specification models the state of a system explicitly, while only an implicit model is provided by the transition-based specification. From this fact, it should not surprise that the *process* involved in producing the specifications encourage quite different strategies. Concretely, the model provided by a state-based specification is explicit in the sense that the values of the state variable are given by a single generator term, without the need for further rewriting. This is not the case with the transition-based specification, where the value of a state variable can be found only by evaluating the corresponding function in the generator term representing the history of the system. This explains the difference in perspective between the specifications. Since in a transition-based specification pairs of state variables and transitions are covered by separate equations, the specifier may concentrate on smaller parts of the system and its specification in the specification process. Consequently, the transition-based approach facilitates a process where the specification of each state variable is done separately from the specification of the other state variables. The mental load in the process is therefore reduced in comparison to the process involved in producing a corresponding state-based specification.

Specification maintenance: One difference of great pragmatic importance relates to the maintenance of specifications. A choice between the state-based or the transition-based specification style may profit from a consideration of what types of modifications are envisaged. By way of illustration, the state-based style may be the most convenient if addition of operations is more likely than addition of state variables. A possible approach is to transform between the two styles of specification as the need arises. With the new HRP Prover, this transformation can be done automatically.

Comprehensibility: Evidently, comprehensibility of specifications is largely a matter of individual judgement. Nevertheless, there are certain aspects of state-based and transition-based specifications that favour one style over the other in this respect. There is also a clear relationship between manual inspection of the specification and animation through execution of the specification:

- **Inspection:** Since the state-based specification “collects” the values of all the state variables into a single generator term, it is easier to see how an operation modifies the system as a whole than with its transition-based counterpart. This perspective is virtually lost in the transition-based specification, where only one state variable is considered in each equation. This is not necessarily problematic if the inspection mainly focuses on pairs of operations and state variables. In this case, the transition-based specification may even be favourable.
- **Animation:** The animation of the state-based specification shows how an operation modifies the system as a whole, by giving the new value of each state variable. This is usually the preferred “mode” of ani-

mation, compared to the transition-based approach where the value of each state variable must be found separately. The transition-based approach may be preferable if the state variables of interest constitute only a small fraction of the total amount of state variables.

Compactness: Since all the state variables are explicitly represented by a single term, a state-based specification is usually more compact than its transition-based counterpart. With n states and m transitions (including initialisations), the specification of the relationship between the states and the transitions is in the standard case given by m equations in a state-based specification. In the corresponding transition-based specification, the same relationship is given by $n \times m$ equations. It is clear that, as the number of states and/or transitions increases, the difference in compactness between a state-based and a transition-based specification becomes startlingly more apparent.

Utilization in other methods: Tools support comprises an important consideration in relation to the translation of state-based and transition-based algebraic specifications into other languages. Since not all tools are supporting both a state-based and a transition-based style of specification, it is important to be able to change between the two types of specifications. When considering various specification languages, we find that one or the other of the two styles of specifications is closer to the "normal practice" adopted when writing specifications in these languages. By way of example, schemas in the Z notation [23] as well as composite objects in VDM specifications [13] adopt the state-based style, while specifications in the Larch Shared Language [9] follows the traditional algebraic approach underlying the transition-based style. Since a state-based specification can be transformed into a transition-based specification, and vice versa, these transformations are useful in transforming certain specifications in one language into specifications in another. This is clearly beneficial to software development projects involving the use of a combination of specification languages.

7 Implications on Software Quality

As is discussed in [22], there are in general two principal ways of ensuring software quality - one in terms of process, and another in terms of product. In a process-oriented approach, quality is seen as an outcome of a good software development process. In a product-oriented approach, quality is assessed or ensured by directly evaluating a given piece of software. Intuitively, it is easy to see that neither the process-oriented nor the product-oriented approach is fully satisfactory. By way of illustration, the *process-oriented approach* may fail due to over-emphasis on traditional quality assurance activities on the cost of adequate testing and verification. On the other hand, the *product-oriented approach* may fail due to insufficient consideration of how the testing and verification can be facilitated by controlling the software development process. These pitfalls immediately suggest that an optimal approach requires a combination of the process-oriented and the product-oriented approach.

In general, formal software development integrates process/product quality by making the evolving software an integral part of the process. It is therefore to be expected that the use of formal methods influences software product quality in very specific ways. There is however a need for empirical results on this influence, which suggests that much effort should be invested into utilizing experiences from applications. It is not necessarily problematic that these applications are strictly limited in terms of the method used and how it is used. On the contrary, it appears that measures on the influence of formal methods need to be rather detailed in order to be really useful in safety assessment.

The development of the new HRP Prover provides insight into how the method influences product quality. In the following, we will summarize some of the findings by suggesting how the six main quality charac-

teristics of ISO 9126 [10] are influenced. This should not be confused with giving an evaluation of the method. The purpose is to identify principles behind the method that call for further work on establishing adequate measures on their actual influence.

Functionality

By functionality, we mean the existence of certain functions and their properties. The functions are those that satisfy stated or implied needs.

- Due to the general concepts of data types and parameterization, the specification language facilitates the construction of a hierarchical well-structured specification. By means of automatic code generation, a corresponding structure is achieved in the implementation while reflecting the specified requirements. As a consequence, requirements can more easily be traced back from code, which greatly improves the efficiency of software maintenance. Improved traceability of requirements also reduces the risk of unintentional changes in functionality during this phase of the software life-cycle.
- Executability of specifications improves the possibility to demonstrate the specified functionality before further development is initiated. By means of automatic code generation, an executable specification can be implemented directly into code exhibiting the same functionality. In this way, questions about the coherence of implementation with respect to a specification to a large extent reduce to a question about the correctness of the code generation algorithm.

Reliability

The reliability of a software system refers to the attributes that bear on its capability to maintain its level of performance under stated conditions for a stated period of time. While specification of these attributes is not part of the method, the level of and confidence in reliability is clearly influenced by several factors of the development process.

- The use of a formal specification language makes it possible to analyse the specification with respect to desired properties, such as safety invariants. In the reported approach, this analysis is carried out in terms of animation (execution of the specification without the need to generate a prototype) and theorem proving. The analysis is supported by the use of the HRP Prover.
- Whether or not automatic code generation is employed, the design can be directed towards so-called safe subsets of programming languages. In the EvalFM project (see section 2), the specification of a reactor safety system was designed into a safe subset of Pascal.
- Automatic code generation from the specification improves the confidence in the ability of the code to correctly reflect the functional requirements.
- Modularity and parameterization facilitates re-use of specifications and code for which the functionality is well known and the reliability has been demonstrated in other systems.

Usability

By usability, we mean attributes that bear on the effort needed for use by a given set of users. It is important that the perspective of the potential users is taken into consideration in individual assessment of the software specification.

- Adequate abstraction level in the specification facilitates the formulation of user requirements.
- Modular, hierarchical specifications facilitate layered specification and implementation of adaptations or extensions, of relevance to the customer or to specific user groups.

Efficiency

The efficiency of a software system relates to the relationship between the level of performance of the software and the amount of resources used, under stated conditions.

- Efficiency of the implementation can be treated as a separate concern in the design of the system.

Maintainability

The maintainability of a software system refers to the effort needed to make specified modifications. We have already seen that there is a clear relationship between maintainability and the well-structuredness and modularity of the specification and the software.

- User requirements can be easily traced back from the given pieces of code.
- New or modified code can be generated directly from the specification of new or changed requirements.
- Identification and implementation of necessary modifications and extensions of the program code is facilitated by a layered, modular implementation reflecting the structure and requirements of the specification.

Portability

Finally, attributes related to portability have a bearing on the ability of the software to be transferred from one environment to another.

- The use of parameterized data types and modular specifications facilitates reuse of specifications and code in new environments.
- The hierarchical structure of the specification facilitates extensions and adaptations of the program.

Evidently, the influence of the use of formal methods on software product quality will depend on the choice of methods and on the way these methods are used. In the reported project, several of the findings above relate to the particular specification method used, in this case algebraic specifications, and the use of e.g. automatic code generation. For other methods or ways of using these, the influences on software product quality will of course vary.

8 Transferability of Results

It is widely accepted that successful use of formal methods requires the availability of reliable support tools. Of this reason, the formal development of a theorem prover is an important undertaking, regardless of the possible benefits findings from the development project may bring to other contexts. Such a restricted view would however fail to recognize the overall aim of the formal development of the HRP Prover, where the transferability of the various techniques and concepts are considered as being of special importance. In particular, the development aims at demonstrating

- the practical use and general relevance of the algebraic specifications in the development of language-oriented tools;
- the impact of formal specification and code generation on various aspects of product quality;
- general principles that can be transferred to other specification languages;
- strategies to combining specification languages.

Since the functionality of the HRP Prover is based on manipulation and analysis of texts in a formal lan-

guage, the specification of the tool involves many concepts and techniques applicable to other language-oriented tools. While many of these concepts and techniques are related to compiler technology, the use of algebraic specifications provides additional advantages related to representation, analysis, transformation, and code generation.

Representation: Basic to these advantages of the approach is the flexibility and naturalness by which abstract syntax is represented in the specification language. Abstract syntax trees correspond closely to the hierarchical structures defined by the generator terms in algebraic specifications. In the formal development of the HRP Prover, the represented texts are algebraic specifications that are translated into a subset of Prolog. The same approach is of course applicable to other source and target languages. By way of example, the theory developed for the translation of Petri nets into algebraic specifications can be used as a basis for extending the HRP Prover with functionality for automating this translation.

Analysis: The hierarchical structure of the abstract syntax trees relates directly to the fact that a single ground algebraic term is sufficient for representing a text in the specified language. This means that analysis can be specified in terms of functions or predicates that take such terms as arguments, where detailed analysis may be distributed to more specialized functions or predicates covering only parts of the abstract syntax. The same general approach is applicable to a many different kinds of analysis, like type correctness of a strongly typed language (static semantics), execution of computer programs (dynamic semantics), characterization of texts in a formal language (e.g. state-based or transition-based specifications), pattern recognition, satisfaction of constraints defining language subsets, etc.

Transformation: The abstract syntax tree representation also simplifies the specification of how texts can be transformed or otherwise manipulated. Examples are refinement techniques, translation of texts between different languages (e.g. translation from Petri nets to algebraic specifications, and from algebraic specifications to Prolog), transformation between different classes of texts within a language (e.g. between state-based and transition-based algebraic specifications), text editing, formatting, and printing, program transformation, etc.

Code generation: Finally, the abstract syntax tree representation benefits the specification of code generation as a special case of transformation. In the formal specification of the HRP Prover, this is illustrated by the translation from the abstract syntax tree representation of algebraic specifications into that of Prolog, followed by the "coding" into concrete Prolog programs.

9 Conclusions

A concrete goal of the formal development of the HRP Prover has been to deliver a new version of the tool, developed in accordance to the methodology supported by the same tool. At a more general level, the development aims at demonstrating the practical use and general relevance of algebraic specifications in the development of language-oriented tools, the impact of formal specification and code generation on various aspects of product quality, general formal development principles that can be transferred to other specification languages, and strategies to combining specification languages.

The formal development of the HRP Prover has delivered a tool that facilitates formal development of modular, well-structured programs by the use of automatic program generation from specification or design. As a consequence, functional requirements can be realized directly, while improving the traceability of requirements from the program code. This also simplifies maintenance and further development, since new program code can be constructed directly from changed or added requirements. The existence of the tool provides evidence to the claim that the methodology based on algebraic specifications and the HRP Prover

can be used efficiently in the development of programs of realistic size and complexity.

The new HRP Prover appears to be the first tool of its kind that has been formally developed in accordance to the same methodology as is supported by the tool. It also appears to be unmatched from earlier developments of theorem provers in its extensive use of algebraic specification. The reported results from the development project are of relevance to the formal development of a wide range of language-oriented tools, involving aspects like analysis, transformation, and code generation. In particular, the approach employed facilitates combination of complementary specification notations.

The core functionality of the new HRP Prover centres around a uniform approach to evaluation and theorem proving in algebraic specifications. An example of a useful extension of the functionality relates to automatic transformation between so-called *state-based* and *transition-based* algebraic specifications. It has been demonstrated how the transformation of specifications between the two classes makes it possible to combine specifications written in different styles. The importance of this novel feature of the new HRP Prover was identified in relation to the work on Petri nets and algebraic specifications. This research activity was initiated in a co-operative project between ENEA and the Halden Project on the combination of graphical and textual notations in formal specification. The overall aim of this project was to contribute to a clarification of the relationship between graphical descriptions and formal specifications, and to provide guidelines for how they can be combined in order to utilize the strengths of each approach. The research activities have continued with the specific aim of generalizing an approach to the translation of Petri nets into algebraic specifications. Petri nets represent in this context an intermediate language between graphical descriptions and algebraic specifications.

The concepts of state-based and transition-based specifications have been applied in the establishment of a uniform approach to the translation of a wide variety of *autonomous* and *non-autonomous* Petri nets into algebraic specification. The approach involves translating Petri nets optionally into state-based or transition-based algebraic specifications, and using automatic transformation between these two classes in order to utilize their relative merits. The translation makes it possible to analyse the nets with techniques established for algebraic specification, including the use of the HRP Prover. Furthermore, the nets can be manipulated indirectly through transformation of the resulting specifications. In this way, Petri nets and algebraic specifications can be integrated to form a common specification document to be used as a basis for the software development.

In order to facilitate more wide-spread use of the new HRP Prover, further work will concentrate on support, maintenance, and dissemination activities. This includes regular software maintenance of the new HRP Prover with its various extensions and support tools, application in software development projects, and based on these experiences, continuous improvement of the instruction material to the tutorial programme. Further work will also concentrate on the development of a suite of tools facilitating the integrated use of Petri nets and algebraic specifications in formal software development. The development of these tools will be carried in accordance to the formal method supported by the HRP Prover, and will be fully integrated as extensions to this tool.

10 References

- [1] L. Beltracchi, NRC research activities. In D.R. Wallace, B.B. Cuthill, L.M. Ippolito and L. Beltracchi (eds.). *Proc. Digital Systems Reliability and Nuclear Safety Workshop*, Sep. 13-14, 1993. NUREG/CP-0136, United States Nuclear Regulatory Commission, Washington DC, 1994.

- [2] J.A. Bergstra, J. Heering and P. Klint (eds.). *Algebraic Specification*. ACM Press, Addison-Wesley, 1989.
- [3] G.H. Chisholm, B.T. Smith, and A.S. Wojcik. Formal specifications for safety grade systems. In *Proc. Methodologies, Tools, and Standards for Cost-Effective, Reliable Software Verification and Validation*. EPRI, CA, USA, Jan. 1992.
- [4] R.H. Crane. Experience gained in the production of licensable safety-critical software for Darlington NGS. In *Proc. Methodologies, Tools, and Standards for Cost-effective, Reliable Software Verification and Validation*. EPRI, Palo Alto, CA, USA, Jan. 1992.
- [5] M.H. van Emden and K. Yukawa. Logic Programming with Equations. *Journal of Logic Programming*, 4: 265-288, 1987.
- [6] European Commission. *European nuclear regulators' current requirements and practices for the licensing of safety critical software for nuclear reactors*. Draft report, revision 8. Luxembourg: Office for Official Publications of the European Communities, 1998.
- [7] K. Futatsugi, J.A. Goguen, J.-P. Jouannaud and J. Meseguer. Principles of OBJ2. In *Proc. ACM Princ. of Prog. Lang.*, pages 52-66, 1985.
- [8] J.V. Guttag. *The Specification and Application to Programming of Abstract Data Types*. Ph.D. Thesis, Computer Science Department, University of Toronto, 1975.
- [9] J.V. Guttag and J.J. Horning. *Larch: Languages and Tools for Formal Specification*. Springer-Verlag, 1993.
- [10] ISO 9126. *International Standards Organisation. Information technology - Software product evaluation - Quality characteristics and guidelines for their use*. ISO/IEC IS 9126.
- [11] N.A. Jaleel and H. Nicholson. *Petri Nets and Fault Diagnosis in Nuclear Reactors*. Research report no. 413, Department of Control Engineering, University of Sheffield, Nov. 1990.
- [12] K. Jensen. An introduction to the theoretical aspects of Coloured Petri Nets. In J.W. de Bakker, W.-P. de Roever, and G. Rozenberg (eds.). *A Decade of Concurrency*, LNCS 803, pages 230-272. Springer-Verlag, 1994.
- [13] C.B. Jones, *Systematic Software Development Using VDM*. Prentice-Hall, 1986.
- [14] J. Knight and B. Littlewood. Critical task of writing dependable software. *IEEE Software*, pages 16-20, Jan. 1994.
- [15] K.H. Mortensen and V. Pinci. Modelling the work flow of a nuclear waste management program. In R. Valette (ed.). *Application and Theory in Petri Nets 1994. Proc. 15th Int'l Petri Net Conference*, LNCS 815, pages 376-395. Springer-Verlag, 1994.
- [16] J.L. Peterson. *Petri Net Theory and the Modeling of Systems*. Prentice-Hall, 1981.
- [17] J. Prock. A 3 level conception for signal validation and early fault detection in closed systems. Enlarged Halden Programme Group Meeting (OECD Halden Reactor Project), Bolkesjø, Norway, 1990.
- [18] J. Prock. A new technique for fault detection using Petri Nets. *Automatica*, 27: 239-245, 1991.
- [19] W. Reisig. *Petri Nets, An Introduction*. Springer-Verlag, 1985.

- [20] T. Sivertsen. Algebraic Specification Used in Formal Software Development - Part 1: Specification, HWR-331, OECD Halden Reactor Project, Norway (1992).
- [21] T. Sivertsen. A case study on the formal development of a reactor safety system. In *Proc. FME'96: Industrial Benefit and Advances in Formal Methods*, Oxford, LNCS 1051, pages 18-38. Springer-Verlag, 1996.
- [22] T. Sivertsen. Integration of software process/product quality. In *Proc. IAEA Specialists' Meeting on Licensing of Backfitting/Modernization of Instrumentation and Control systems Important to Safety*. Vienna, Austria, September 7-9, 1998.
- [23] J.M. Spivey. *The Z Notation - A Reference Manual*. Prentice-Hall, 1989.
- [24] W. Swartout and R. Balzer. On the inevitable intertwining of specification and implementation. *Comm. ACM*, 25(7): 438-440, 1982.
- [25] D.R. Wallace, B.B. Cuthill, L.M. Ippolito and L. Beltracchi (eds.). *Proc. Digital Systems Reliability and Nuclear Safety Workshop*, Sep. 13-14, 1993. NUREG/CP-0136, United States Nuclear Regulatory Commission, Washington DC, 1994.

REEVALUATION OF REGULATORY GUIDANCE ON MODAL RESPONSE COMBINATION METHODS FOR SEISMIC RESPONSE SPECTRUM ANALYSIS

R.J. Morante and Y.K. Wang
Brookhaven National Laboratory
Upton, New York 11973
and
W.E. Norris
U.S. Nuclear Regulatory Commission

ABSTRACT

Regulatory Guide 1.92 "Combining Modal Responses and Spatial Components in Seismic Response Analysis" was last revised in 1976. The purpose of this project was to re-evaluate the current regulatory guidance for combining modal responses in response spectrum analysis, evaluate recent technical developments, and recommend revisions to the regulatory guidance. In addition, Standard Review Plan (SRP) Section 3.7.2, "Seismic System Analysis," was also reviewed to identify related sections which may need to be revised. The objectives were addressed through an evaluation of past studies, supplemented by analysis of a piping system model previously utilized in NUREG/CR-5627, "Alternate Modal Combination Methods in Response Spectrum Analysis".

The project evaluated (1) methods for separation of the in-phase and out-of-phase modal response components; (2) methods for combination of the out-of-phase modal response components; (3) the contribution of "missing mass"; and (4) combination of the three elements of response to produce the total response. Numerical results from response spectrum analysis were compared to corresponding time history analysis results to assess the accuracy of the various combination methods tested.

The methods selected for evaluation were those which have been subjected to the greatest level of prior review and assessment. For separation of in-phase from out-of-phase modal response components, the methods proposed by Lindley-Yow, Hadjian and Gupta were evaluated. For combination of the out-of-phase modal response components, the Square Root of the Sum of the Squares (SRSS), NRC Grouping, NRC Ten Percent, NRC-Double Sum Combination (DSC), Rosenblueth's DSC, and Der Kiureghian's Complete Quadratic Combination (CQC) methods were evaluated. For treatment of the "missing mass" contribution, the method of Kennedy was evaluated. Response spectrum analyses were conducted by combining elements of the above methods to construct complete response spectrum analysis solutions.

Based on the qualitative evaluation and the numerical results generated in this project, it was concluded that Rosenblueth's DSC and Der Kiureghian's CQC methods for combining out-of-

phase modal response components, coupled with the Lindley-Yow or Gupta method for separation of in-phase from out-of-phase modal response components, and inclusion of Kennedy's missing mass contribution to account for modes with frequencies above the Zero Period Acceleration (ZPA) frequency of the response spectrum constitute the best methodologies currently available for response spectrum analysis.

The NRC Grouping, Ten Percent and DSC methods for combining out-of-phase modal response components produced more conservative but less accurate results; removal of these methods from the Regulatory Guide should be considered, because absolute summation of closely spaced modal responses has no documented technical basis. SRSS remains applicable in the absence of closely spaced nodes.

Separation of modal responses into in-phase and out-of-phase components for modes with frequencies below the ZPA frequency of the response spectrum produced more accurate results than commonly applied past methods, in which all modes below the ZPA frequency were considered to be out-of-phase. It is important to note that at low frequency ($<$ the frequency of the peak spectral acceleration), modal responses should be treated as out-of-phase. A limitation of the Lindley-Yow formulation is that low frequency modal responses are separated into out-of-phase and in-phase components; consequently, when significant low frequency modal responses exist, the Lindley-Yow formulation must be appropriately modified. The Gupta formulation correctly assumes low frequency modes are out-of-phase. This project did not evaluate mode response combination methods applied to systems with significant low frequency response.

The present paper describes the qualitative evaluation of modal response combination methods. The numerical results, detailed conclusions, and specific recommendations for revision of regulatory guidance are documented in a NUREG/CR, scheduled for publication in Spring 1999.

1 INTRODUCTION

General Design Criterion 2, "Design Basis for Protection Against Natural Phenomena" of Appendix A, "General Criteria for Nuclear Power Plants, to 10CFR Part 50, "Licensing of Production and Utilization Facilities" specifies a requirement that nuclear power plant structures, systems, and components which are important to safety be designed to withstand the effects of earthquakes, without loss of capability to perform their safety functions. In addition, Paragraph (a)(1) of Section VI of Appendix A to 10CFR Part 100, "Reactor Site Criteria" identifies the use of a suitable dynamic analysis as one method of ensuring that structures, systems, and components can withstand seismic loads.

The United States Nuclear Regulatory Commission (NRC) issues Regulatory Guides (RG) which describe methods acceptable to the NRC staff for satisfying regulations. One such guide is RG 1.92, "Combining Modal Responses and Spatial Components in Seismic Response Analysis," (Reference 1). This guide was last revised in 1976, prior to a number of significant technical developments for combining modal responses.

The 1989 revision to Standard Review Plan (SRP) Section 3.7.2, "Seismic System Analysis," (Reference 2) recognized a number of recent technical developments by reference, and stated that their application to nuclear power plant seismic analysis is subject to review on a case-by-case basis. Also incorporated into SRP Section 3.7.2 as Appendix A was a procedure to address high frequency mode effects, developed by Kennedy (Reference 3).

The objective of this project was to evaluate these recent developments for modal response combination, through a literature review and analytical effort, and to provide recommendations for revision of RG 1.92 and SRP Section 3.7.2 which reflect the current state of technology for combining modal responses in seismic response spectrum analysis..

The design of structures, systems, and components for seismic loads is complicated by the uncertainty about the future seismic event. For seismic analysis where an accurate record of the seismic input exists, time history analysis is the best approach to mathematical prediction of seismic structural response. In the seismic design process, two primary approaches are available to account for the uncertainty in the seismic ground motion: perform a number of time history analyses utilizing an appropriate set of acceleration records or perform response spectrum analysis utilizing a bounding peak acceleration vs. frequency spectrum. Either approach, properly implemented, should provide a conservative basis for seismic design.

A significant feature of response spectrum analysis is that only the maximum dynamic response is predicted. In time history analysis, the response vs. time is predicted. The mathematical simplicity of response spectrum analysis is achieved by calculating the independent peak modal responses and then applying a rule for combining these, in order to predict the peak dynamic response. The input response spectrum defines the acceleration to be applied to each natural mode of vibration of the structure, depending on its modal frequency. However, the spectrum provides no information about the time phasing of excitation of these modes.

It is evident that the modal response combination method used to predict the peak dynamic response is a critical element of the response spectrum analysis method. Many researchers have studied this specifically for seismic analysis, and a number of increasingly sophisticated methods have been developed. In 1984, Kennedy (Reference 3) reviewed alternate methods for modal response combination for the NRC and provided recommendations for revision to regulatory guidance. Gupta (Reference 4, Chapter 3) provides an excellent review of these developments up to the late 1980's, including his own modal response combination method. No significant technical developments emerged in the 1990's.

This present investigation does not address the issue of seismic design methodology - time history analysis vs. response spectrum analysis. It is focused on comparison and evaluation of different modal response combination methods for use in response spectrum analysis. Because the objective of response spectrum analysis is to predict, with reasonable accuracy, the peak dynamic response to a time varying acceleration input, comparison to time history solutions is the primary method employed to evaluate the applicability and limitations of the various modal response combination methods.

The initial phase of this program focused on review of the technical literature and selection of candidate modal response combination methods for more detailed evaluation. Acceptable methods in RG 1.92 were also included to provide a comparison to more recent technical developments. References 3 and 4 provided an excellent starting point. In addition, prior numerical studies conducted by Brookhaven National Laboratory (References 5, 6, 7) on response spectrum analysis of piping systems provided a quantitative database for alternate modal response combination methods. Industry standards such as ASCE Standard 4 (References 8, 9) were also reviewed. The work of Maison et al (Reference 10) provided a numerical study of different modal response combination methods applied to a building structure. A significant body of additional reference material was included in the literature review.

The methods selected for evaluation were those which have been subjected to the greatest level of prior review and assessment. The evaluation addressed (1) methods for combination of the out-of-phase modal response components; (2) methods for separation of the in-phase and out-of-phase modal response components; (3) the contribution of high frequency modal responses; and (4) combination of the three elements of response to produce the total response. The term "in-phase" denotes response that is in-phase with the time varying acceleration input; the term "out-of-phase" denotes response that is out-of-phase with the time varying acceleration input. Gupta (Reference 4) refers to these as the "rigid" or "pseudo-static" response and the "damped periodic" response, respectively.

It is important to note that individual modal responses which are each "out-of-phase" with respect to the time varying acceleration input may be nearly in-phase with each other. This is commonly referred to as the "closely spaced modes" issue, because modes close in frequency are considered most likely to respond nearly in-phase when excited by the same time varying acceleration input. This is addressed by methods for combination of the "out-of-phase" modal responses.

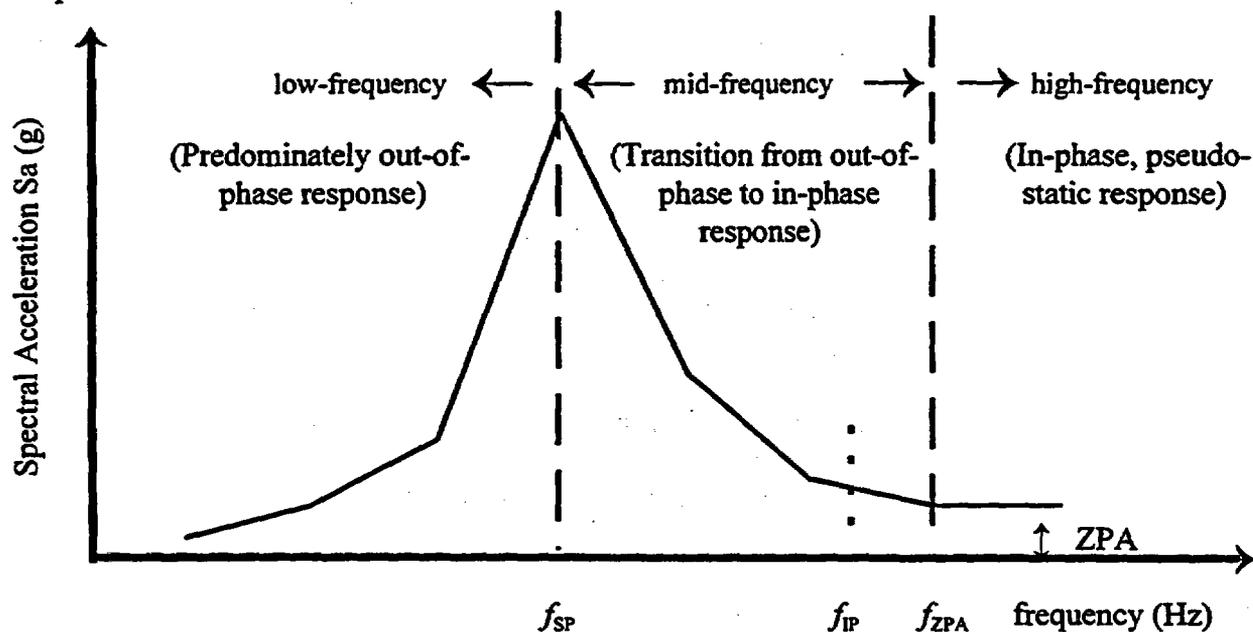
Modes with frequencies higher than the frequency (f_{ZPA}) at which the spectral acceleration returns to the Zero Period Acceleration (ZPA) respond pseudo-statically, in-phase with the time varying acceleration input and, therefore, in-phase with each other. The contribution of these modes to the total response is most accurately and efficiently treated by static analysis of the "missing mass" (i.e., system mass not participating in the modes with frequencies below f_{ZPA}) multiplied by the ZPA. Modes with frequencies in the amplified region of the response spectrum ($f < f_{ZPA}$) generally have two components of response: an out-of-phase component and an in-phase component. The in-phase modal response components and the missing mass contribution are combined algebraically, to produce the total in-phase response component.

For combination of the out-of-phase modal response components, Square Root of the Sum of the Squares (SRSS), NRC Grouping, NRC Ten Percent, NRC-DSC, Rosenblueth's DSC (Reference 11), and Der Kiureghian's CQC (Reference 12) methods were evaluated. For separation of in-phase from out-of-phase modal response components, the methods proposed by Lindley and Yow (Reference 13), Hadjian (Reference 14), and Gupta (Reference 4) were evaluated. For treatment of the missing mass contribution, the method of Kennedy (Reference 3) was evaluated.

2 DESCRIPTION OF MODAL RESPONSE COMBINATION METHODS

To lay the groundwork for the ensuing discussions of modal response combination methods, it is necessary to define the different regions of a typical seismic response spectrum and key frequencies which divide these regions. The major application of seismic response spectrum analysis is for systems and components attached to building structures.

A building-filtered in-structure response spectrum depicting spectral acceleration vs. frequency is the typical form of seismic input for such analyses. This type of spectrum usually exhibits a sharp peak at the fundamental frequency of the building/soil dynamic system. A typical, idealized in-structure response spectrum is shown below; the spectral regions and key frequencies are indicated.



f_{SP} \equiv frequency at which the peak spectral acceleration is reached; typically the fundamental frequency of the building/soil system

f_{ZPA} \equiv frequency at which the spectral acceleration returns to the zero period acceleration (ZPA)

f_{IP} \equiv frequency above which the SDOF modal responses are in-phase with the time varying acceleration input used to generate the spectrum

The high frequency region of the spectrum ($> f_{ZPA}$) is characterized by no amplification of the peak acceleration of the input time history. A SDOF oscillator having a frequency $> f_{ZPA}$ is accelerated in-phase and with the same acceleration magnitude as the applied acceleration, at each instant in time. A system or component with fundamental frequency $> f_{ZPA}$ is correctly

analyzed as a static problem subject to a loading equal to mass times ZPA. The system or component is said to respond "pseudo-statically." This concept can be extended to the high frequency ($> f_{ZPA}$) modal responses of multi-modal systems or components. The mass not participating in the amplified modal responses (i.e., "missing mass") multiplied by the ZPA is applied in a static analysis, to obtain the response contribution from all modes with frequencies $> f_{ZPA}$.

In the low-frequency region of the spectrum ($< f_{SP}$) the modal responses of SDOF oscillators are not in-phase with the applied acceleration time history, and generally are not in-phase with each other. These are designated "out-of-phase" modal responses. Since a response spectrum provides only peak acceleration vs. frequency, with no phasing information, the out-of-phase peak modal responses for a multi-modal structural system requires a rule or methodology for combination. Based on the assumption that the peak modal responses are randomly phased, the square root of the sum of the squares (SRSS) method was developed and adopted. Modifications to SRSS were subsequently developed, in order to account for potential phase correlation when modal frequencies are numerically close (i.e., "closely spaced modes").

In the mid-frequency region (f_{SP} to f_{ZPA}), it has been postulated that the peak SDOF oscillator modal responses consist of two distinct and separable elements. The first element is the out-of-phase response component and the second element is the in-phase response component. It is further postulated that there is a continuous transition from out-of-phase response to in-phase response. If $f_{IP} < f_{ZPA}$ can be defined, then the mid-frequency region can be further divided into two sub-regions: $f_{SP} < f < f_{IP}$ and $f_{IP} \leq f \leq f_{ZPA}$.

It is noted that past practice in the nuclear power industry has been to assume that individual modal responses in the mid-frequency region ($f_{SP} < f < f_{ZPA}$) are out-of-phase, and that the combination methods applicable to the low-frequency region are also applicable to the mid-frequency region.

Three elements are needed to define a suitable methodology for the mid-frequency region:

- 1) A definition for f_{IP} .
- 2) A method for separating the in-phase and out-of-phase components of individual peak modal responses.
- 3) A phase relationship for combining the total out-of-phase response component with the total in-phase response component.

Methods for modal response combination will be described in the following sections:

- Section 2.1 - Combination of Out-of-Phase Modal Response Components
- Section 2.2 - Separation of Out-of-Phase and In-Phase Response Components in Mid-Frequency Region
- Section 2.3 - Contribution of High-Frequency Modes
- Section 2.4 - Complete Solution for Response Spectrum Analysis

Terms used in the following sections are:

- Sa_i = Spectral Acceleration for mode i
- R_i = Response of mode i
- α_i = In-phase response ratio for mode i
- Rr_i = In-phase response component for mode i
- Rp_i = Out-of-phase response component for mode i
- Rr = Total in-phase response component from all modes
- Rp = Total out-of-phase response component from all modes
- Rt = Total combined response from all modes
- C_{jk} = Modal response correlation coefficient between modes j and k.

2.1 Combination of Out-of-Phase Modal Responses Components

In generalized form, all of the out-of-phase modal response combination methods can be represented by a single equation:

$$R_p = \left[\sum_{j=1}^n \sum_{k=1}^n C_{jk} R_{p_j} R_{p_k} \right]^{1/2} \quad (\text{Eqn. 2 - 1})$$

The coefficients C_{jk} can be uniquely defined for each method.

2.1.1 Square Root of the Sum of the Squares (SRSS)

At the foundation of all methods for combining uncorrelated modal responses is SRSS. All of the methods for combination of the out-of-phase response components are equivalent to SRSS if there are no "closely spaced" modes.

In the case of SRSS,

$$\begin{aligned} C_{jk} &= 1.0 \text{ for } j = k \\ C_{jk} &= 0.0 \text{ for } j \neq k \end{aligned} \quad (\text{Eqn. 2 - 2})$$

SRSS Combination reduces to:

$$R_p = \left[\sum_{i=1}^n R_{p_i}^2 \right]^{1/2} \quad (\text{Eqn. 2 - 3})$$

2.1.2 NRC Grouping Method

The NRC Grouping Method (Reference 1) is the most commonly applied method of accounting for closely spaced modes in the nuclear power industry. The system modal responses are grouped and summed absolutely before performing SRSS combination of the groups. To illustrate the method, consider a system with 10 modes R_{p_1} through $R_{p_{10}}$ and associated frequencies f_1 through f_{10} with the following distribution:

$$\begin{array}{rcl}
 f_2 & > & 1.1 f_1 \\
 f_3, f_4 & < & 1.1 f_2 \\
 f_5 & > & 1.1 f_2 \\
 f_6 & < & 1.1 f_5 \\
 f_7 & > & 1.1 f_5 \\
 f_8 & > & 1.1 f_7 \\
 f_9, f_{10} & < & 1.1 f_8
 \end{array}$$

The modal responses are grouped such that the lowest and highest frequency modes in a group are within 10% and no mode is in more than one group. Using the distribution above, the following groups are created:

$$\begin{aligned}
 GR_1 &= R_{p_1} \\
 GR_2 &= |R_{p_2}| + |R_{p_3}| + |R_{p_4}| \\
 GR_3 &= |R_{p_5}| + |R_{p_6}| \\
 GR_4 &= R_{p_7} \\
 GR_5 &= |R_{p_8}| + |R_{p_9}| + |R_{p_{10}}|
 \end{aligned}$$

$$R_p = \left[\sum_{i=1}^n GR_i^2 \right]^{1/2} \quad (\text{Eqn. 2 - 4})$$

In this illustration, $n = 5$.

The major criticism of the NRC Grouping Method is the use of absolute summation within each group. If modal responses are assumed to be correlated because they have closely spaced frequencies, then summation should be algebraic within each group. The bias toward conservatism in the NRC Grouping Method is somewhat contradictory to the basic premise for grouping.

Definition of $C_{j,k}$ for the NRC Grouping Method is somewhat cumbersome:

$$\begin{array}{ll}
 C_{j,k} = 1.0 & \text{for } j = k \\
 C_{j,k} = 0.0 & \text{for } j \neq k, \text{ not in same group} \\
 C_{j,k} = 1.0 & \text{for } j \neq k, \text{ in the same group, } R_{p_j} \text{ and } R_{p_k} \text{ have same sign} \\
 C_{j,k} = -1.0 & \text{for } j \neq k, \text{ in the same group, } R_{p_j} \text{ and } R_{p_k} \text{ have opposite sign}
 \end{array} \quad (\text{Eqn. 2 - 5})$$

In implementing the NRC Grouping Method, the approach presented in the illustration is more straightforward.

2.1.3 NRC Ten Percent Method

The NRC Ten Percent Method (Reference 1) is a generally more conservative variation of the NRC Grouping Method. Closely spaced modes are defined as modes with frequencies within 10% of each other and absolute summation of the closely spaced modal responses is specified. The difference is that modal responses are not grouped.

In terms of the coefficients, C_{jk} , the NRC Ten Percent Method can be defined as follows:

$$\begin{aligned}
 C_{jk} &= 1.0 && \text{for } j = k \\
 C_{jk} &= 0.0 && \text{for } j \neq k, \text{ and } f_j \text{ and } f_k \text{ separated by } > 10\% \text{ of the} \\
 &&& \text{lower frequency} \\
 C_{jk} &= 1.0 && \text{for } j \neq k, \text{ and } f_j \text{ and } f_k \text{ separated by } \leq 10\% \text{ of the} \quad (\text{Eqn. 2 - 6}) \\
 &&& \text{lower frequency; } R_{pj} \text{ and } R_{pk} \text{ same sign} \\
 C_{jk} &= -1.0 && \text{for } j \neq k, \text{ and } f_j \text{ and } f_k \text{ separated by } \leq 10\% \text{ of the} \\
 &&& \text{lower frequency; } R_{pj} \text{ and } R_{pk} \text{ opposite sign}
 \end{aligned}$$

The definition of C_{jk} is analogous to that for the NRC Grouping Method, except that grouping is not performed.

As an illustration of the difference between these two methods, assume three modal responses R_{p1} , R_{p2} , R_{p3} with frequencies f_1 , f_2 , f_3 and a frequency distribution defined as follows:

$$\begin{aligned}
 f_2 &\leq 1.1 f_1 \\
 1.1 f_1 &< f_3 \leq 1.1 f_2
 \end{aligned}$$

By the NRC Grouping Method,

$$R_p = \left(\sum_{i=1}^2 GR_i^2 \right)^{1/2}$$

where

$$\begin{aligned}
 Gr_1 &= |R_{p1}| + |R_{p2}| \\
 Gr_2 &= R_{p3}
 \end{aligned}$$

or

$$R_p = \left(\sum_{i=1}^3 R_{pi}^2 + 2 |R_{p1} * R_{p2}| \right)^{1/2}$$

By the NRC Ten Percent Method,

$$R_p = \left(\sum_{i=1}^3 R_{p_i}^2 + 2 |R_{p_1} * R_{p_2}| + 2 |R_{p_2} * R_{p_3}| \right)^{1/2}$$

The NRC Ten Percent Method has an additional contribution to R_p because $f_3 \leq 1.1 f_2$. The NRC Ten percent Method will always produce results \geq NRC Grouping Method.

2.1.4 NRC Double Sum Combination (NRC-DSC)

The NRC-DSC Method (Reference 1) is an adaptation of Rosenblueth's method, described in Section 2.1.5. The coefficients C_{jk} are defined by Equation 2-7. A conservative modification, consistent with the NRC Grouping and Ten Percent methods, is that the product $C_{jk} R_{p_j} R_{p_k}$ is always taken as positive. In Rosenblueth's method, the product may be either positive or negative, depending on the signs of R_{p_j} and R_{p_k} . Consequently, NRC-DSC will always produce results \geq Rosenblueth's method.

2.1.5 Rosenblueth's Double Sum Combination (DSC)

Rosenblueth (Reference 11) provided the first significant mathematical approach to evaluation of modal correlation for seismic response spectrum analysis. It is based on the application of random vibration theory, utilizing a finite duration of white noise to represent seismic loading. A formula for calculation of the coefficients C_{jk} as a function of the modal circular frequencies (ω_j , ω_k), modal damping ratios (β_j , β_k), and the time duration of strong earthquake motion (t_D) was derived. Using the form of the equation from Reference 1,

$$C_{jk} = \frac{1}{1 + \frac{(\omega_j' - \omega_k')^2}{(\beta_j' \omega_j + \beta_k' \omega_k)^2}}$$

(Eqn. 2 - 7)

where $\omega_{()}' = \omega_{()} [1 - \beta_{()}^2]^{1/2}$

$$\beta_{()}' = \beta_{()} + \frac{2}{t_D \omega_{()}}$$

Numerical values of C_{jk} were tabulated for the DSC Method as a function of frequency, frequency ratio, and strong motion duration time for constant modal damping of 1%, 2%, 5% and 10%. The effect of t_D is most significant at 1% damping and low frequency. For 5% and 10% damping, $t_D = 10$ sec. and 1000 sec. produced similar values for C_{jk} regardless of frequency. The most significant result is that C_{jk} is highly dependent on the damping ratio. For 2%, 5% and 10% damping, $C_{jk} \approx 0.2, 0.5$ and 0.8 respectively, at a frequency ratio of 0.9 (modal frequencies within 10%). In comparison, the definition of closely-spaced modes used in the NRC Grouping and Ten Percent Methods are not damping-dependent.

2.1.6 Der Kiureghian's Complete Quadratic Combination (CQC)

Der Kiureghian (Reference 12) presents a methodology similar to Rosenblueth's Double Sum Combination (Reference 11) for evaluation of modal correlation for seismic response spectrum analysis. It is also based on application of random vibration theory, but utilizes an infinite duration of white noise to represent seismic loading. A formula for calculation of the coefficients C_{jk} as a function of modal circular frequencies and modal damping ratios was derived:

$$C_{jk} = \frac{8 (\beta_j \beta_k \omega_j \omega_k)^{1/2} * (\beta_j \omega_j + \beta_k \omega_k) * \omega_j \omega_k}{(\omega_j^2 - \omega_k^2)^2 + 4 \beta_j \beta_k \omega_j \omega_k (\omega_j^2 + \omega_k^2) + 4(\beta_j^2 + \beta_k^2) \omega_j^2 \omega_k^2} \quad (\text{Eqn. 2 - 8})$$

While the form of Equation 2-8 differs significantly from Equation 2-7, the two equations produce equivalent results if t_D is assumed very large in Equation 2-7.

2.1.7 ASCE Standard 4 Recommended Methods

For combination of out-of-phase modal response components, ASCE Standard 4 (Reference 8) specifies the DSC Method (Equation 3200-16). The NRC methods and CQC are also recognized in the commentary.

Draft ASCE Standard 4 (Reference 9) specifies a modified DSC Method (Equation 3200-19) or the CQC Method (Equation 3200-22) as an alternative. The commentary to the Draft ASCE Standard (Reference 9) indicates that Equation 3200-19 produces correlation coefficients "which are practically the same" as Equation 3200-22. Although not indicated, the modified DSC Method presented in Equation 3200-19 was developed by Gupta (Reference 4).

2.2 Separation of Modal Responses into Out-of-Phase Components and In-Phase Components

Three methods have received considerable prior review and evaluation: Lindley-Yow (Reference 13), Hadjian (Reference 14), and Gupta (Reference 4). It should be noted that the mathematical statement of each method is not restricted to the mid-frequency range ($f_{SP} < f < f_{ZPA}$) of the response spectrum. However, as discussed at the beginning of this chapter, it is in the mid-frequency range that the separation of individual peak modal responses into out-of-phase and in-phase modal response components is applicable. For $f \leq f_{SP}$, modal responses should be considered out-of-phase and combined by the methods presented in Section 2.1. For $f \geq f_{ZPA}$, modal responses are in-phase and are most accurately accounted for by the method of Kennedy (Reference 3).

The similarities and differences, as well as the limitations, of the three methods are described in the following sections.

2.2.1 Lindley-Yow Method

Mathematically, the Lindley-Yow method (Reference 13) is defined by the following equations:

$$\alpha_i = ZPA/Sa_i \quad 0 \leq \alpha_i \leq 1.0 \quad (\text{Eqn. 2 - 9})$$

$$Rr_i = R_i * \alpha_i \quad (\text{Eqn. 2 - 10})$$

$$Rp_i = R_i * \sqrt{1 - \alpha_i^2} \quad (\text{Eqn. 2 - 11})$$

$$Rr = \sum_{i=1}^n Rr_i \quad (\text{Eqn. 2 - 12})$$

$$Rp = \left[\sum_{j=1}^n \sum_{k=1}^n C_{jk} Rp_j Rp_k \right]^{1/2} \quad (\text{Eqn. 2 - 13})$$

$$Rt = \sqrt{Rr^2 + Rp^2} \quad (\text{Eqn. 2 - 14})$$

where the C_{jk} 's are determined by the selected method for combining the out-of-phase modal response components described in Section 2.1.

From these mathematical relationships, the following characteristics of the Lindley-Yow method are observed:

- $\alpha_i \rightarrow 1.0$ as $f_i \rightarrow f_{ZPA}$ ($Sa_i = ZPA$). Consequently, $f_{IP} = f_{ZPA}$ in the Lindley-Yow method.
- The in-phase component of modal response for every mode has an associated acceleration equal to the ZPA.
- The out-of-phase component of an individual peak modal response has an associated modified spectral acceleration given by

$$\bar{S}a_i = \left(Sa_i^2 - ZPA^2 \right)^{1/2} \quad (\text{Eqn. 2 - 15})$$

- $R_i = (Rp_i^2 + Rr_i^2)^{1/2}$; which infers that the in-phase and out-of-phase response components of an individual peak modal response are uncorrelated and, therefore, combine by SRSS.
- All in-phase modal response components (Rr_i) are summed algebraically to obtain Rr .
- All-out-of-phase modal response components (Rp_i) are combined by a suitable method (as described in Section 2.1) to obtain Rp .
- The total response, Rt , is obtained by SRSS combination of Rr and Rp ; i.e., Rr and Rp are uncorrelated.
- α_i attains its minimum value at $f_i = f_{SP}$, but increases for $f_i < f_{SP}$ until it attains a value of 1.0 when $Sa_i = ZPA$ in the low frequency region of the spectrum. Values of $\alpha_i > 1.0$ have no meaning because $(1 - \alpha_i^2)^{1/2}$ becomes imaginary.

An obvious limitation of the Lindley-Yow method is in the low frequency range ($f < f_{SP}$) of the response spectrum. There is no physical basis for assuming that low frequency modal responses become increasingly in-phase with the input acceleration time-history, which is an outcome if the Lindley-Yow method is applied to low frequency modal responses. Modal responses in the low frequency range are generally out-of-phase with the input acceleration time history. Therefore, the Lindley-Yow method is applicable to structural systems which do not have significant modal responses with $f_i < f_{SP}$. Lindley and Yow (Reference 13) do not address this limitation. For the sample problems presented in Reference 13, the lowest system frequency is greater than f_{SP} of the applied response spectrum. Therefore, the results reported in Reference 13 are not affected by this limitation. Circumventing this limitation in the Lindley-Yow method is straightforward: apply it only to those modes with $f_i \geq f_{SP}$ and set $\alpha_i = 0$ for $f_i < f_{SP}$.

For a structural system with fundamental frequency $\geq f_{SP}$, the Lindley-Yow method lends itself to a relatively straightforward physical interpretation. In the limit, if all modes are retained in the solution, the total mass participation is unity. Applying the Lindley-Yow method is equivalent to performing a static analysis of the system loaded by total mass times the ZPA, and performing the response spectrum analysis for amplified modes $f < f_{ZPA}$ using modified spectral accelerations, $\bar{S}a_i$ given by Equation 2-15. The total dynamic response is then obtained by SRSS combination.

The Lindley-Yow method automatically provides for algebraic combination of modal responses above f_{ZPA} , since, $\alpha_i = 1.0$; $R_{p_i} = 0$ and $R_{r_i} = R_i$. However, to completely account for the modal response above f_{ZPA} , all system modes of vibration need to be included in the analysis. This contribution is most accurately and efficiently calculated by use of the missing mass method discussed in Section 2.3. Therefore, while in theory the Lindley-Yow method includes the in-phase contribution from modes above f_{ZPA} , its practical application is for modal responses below f_{ZPA} , coupled with the missing mass method for modal contributions above f_{ZPA} . It is noted that the Lindley-Yow/missing mass approach will produce identical results for any modal analysis cutoff frequency $\geq f_{ZPA}$.

2.2.2 Hadjian Method

The Hadjian Method (Reference 14) is similar in formulation to the Lindley-Yow method, with two notable differences:

- Equation 2-11 is replaced by

$$R_{p_i} = R_i * (1 - \alpha_i) \quad (\text{Eqn. 2 - 16})$$

- Equation 2-14 is replaced by

$$R_t = |R_p| + |R_r| \quad (\text{Eqn. 2 - 17})$$

- The modified spectral acceleration is given by

$$\bar{S}a_i = Sa_i - ZPA \quad (\text{Eqn. 2 - 18})$$

The Hadjian method has the same limitation as the Lindley-Yow method in the low frequency range, because the definition of α_i is identical. However, the Hadjian Method possesses internal contradictions with respect to the assumed phase relationships between in-phase and out-of-phase response components. Combining Equations 2-10 and 2-16 yields

$$R_i = R_{p_i} + R_{r_i} \quad (\text{Eqn. 2 - 19})$$

This implies that the in-phase and out-of-phase response components for each mode are in-phase with each other. However, all R_{r_i} 's are in-phase and summed algebraically, per Eqn. 2-12, to obtain R_r . Therefore, it would follow that all R_{p_i} 's are also in-phase and should be summed algebraically to obtain R_p . This contradicts Equation 2-13, in which the R_{p_i} 's are assumed to be predominantly out-of-phase. Kennedy (Reference 3) previously identified this contradiction. On this basis, the Hadjian method is not recommended and was not included in subsequent numerical studies.

2.2.3 Gupta Method

The Gupta Method (Reference 4) is identical in form to the Lindley-Yow method. The one very significant difference is the definition of α_i . Equations 2-10 through 2-14 remain the same. In the Gupta method, α_i is an explicit function of frequency. The original basis for definition of α_i is semi-empirical, derived from numerical studies using actual ground motion records. A best fit equation, which defines α_i as a continuous function of frequency, was developed from the results of the numerical studies.

Two spectrum-dependent frequencies (f_1, f_2) are first defined as follows:

$$f_1 = \frac{S_{a_{\max}}}{2\pi S_{v_{\max}}} \quad (\text{Eqn. 2-20})$$

where $S_{a_{\max}}$ and $S_{v_{\max}}$ are the maximum spectral acceleration and velocity, respectively.

$$f_2 = (f_1 + 2 f_{ZPA})/3 \quad (\text{Eqn. 2-21})$$

Gupta's definition of α_i is given by:

$$\begin{aligned} \alpha_i &= 0 \text{ for } f_i \leq f_1 \\ \alpha_i &= \frac{\ln(f_i/f_1)}{\ln(f_2/f_1)} \text{ for } f_1 \leq f_i \leq f_2 \\ \alpha_i &= 1.0 \text{ for } f_i \geq f_2 \end{aligned} \quad (\text{Eqn. 2-22})$$

For a sharply peaked, in-structure response spectrum,

$$f_1 = f_{SP}$$

because $Sv_{max} = \text{Max} (Sa_i / \omega_i) = Sa_{max} / \omega_{SP}$

Substitution into Equation 2-20 yields

$$f_1 = \frac{\omega_{SP}}{2\pi} = f_{SP}$$

The corresponding definition of f_2 yields

$$f_2 = (f_{SP} + 2f_{ZPA}) / 3$$

For a sharply peaked, in-structure response spectrum, the Gupta method has the following characteristics:

- For $f_i \leq f_{SP}$, $\alpha_i = 0$.
Consequently, all modal responses with $f_i \leq f_{SP}$ are treated as out-of-phase. The limitation in the Lindley-Yow definition of α_i for $f_i \leq f_{SP}$ does not apply to Gupta's method.
- For $f_2 \leq f_i \leq f_{ZPA}$, $\alpha_i = 1.0$
Consequently, all modal responses with $f_i \geq f_2$ are treated as in-phase. This infers that $f_{IP} = f_2$ in the Gupta method.
- Only modal responses with $f_{SP} < f_i < f_2$ are separated into out-of-phase and in-phase response components.

The potential limitations of the Gupta method lie in the semi-empirical basis for definition of α_i as a function of f_i . The range of applicability is difficult to assess without a comprehensive numerical study using ground and in-structure acceleration records. In Reference 4, Gupta indicates that α_i can be numerically evaluated if the time history used to generate the response spectrum is known. It is implied without stating that numerical evaluation of α_i is more accurate than the semi-empirical definition of α_i given by Equation 2-22.

The overall structure of the Gupta method is superior to the Lindley-Yow method because there is no limitation for modal responses with $f_i < f_{SP}$. In addition, any value of $f_{IP} \leq f_{ZPA}$ can be accommodated by setting $f_2 = f_{IP}$, in lieu of Equation 2-21.

For initial numerical studies, the Lindley-Yow method was selected to evaluate the importance of separating modal responses into out-of-phase and in-phase response components. For follow-up numerical studies, the Gupta method was selected in order to evaluate the influence of f_{IP} on the

response spectrum solution. This was accomplished by selecting three (3) different numerical values for f_2 .

2.2.4 ASCE Standard 4 Recommended Methods

For separation of in-phase and out-of-phase response components, ASCE Standard 4 (Reference 8) recognizes the Lindley-Yow, Hadjian, and Gupta Methods in the commentary.

Draft ASCE Standard 4 (Reference 9) specifies separation of the in-phase and out-of-phase response components consistent with Gupta's method (Eqns. 3200-18, 3200-20, and 3200-21) except that $f_2 = f_r$ (defined as the "cutoff frequency or ZPA frequency") is substituted for Eqn. 2-21. The frequency f_r is not clearly defined, but is $\leq f_{ZPA}$. The Lindley-Yow and Hadjian methods are recognized in the commentary to Reference 9.

2.3 Contribution of High Frequency Modes

2.3.1 Missing Mass Method

The "Missing Mass" Method is a convenient, computationally efficient and accurate method to (1) account for the contribution of all modes with frequencies above the frequency (f_{ZPA}) at which the response spectrum returns to the Zero Period Acceleration (ZPA) and (2) account for the contribution to support reactions of mass which is apportioned to system support points. It constitutes the total effect of all system mass which does not participate in (i.e., "missing" from) the modes with frequencies below f_{ZPA} . The system response to the missing mass is calculated by performing a static analysis for applied loads equal to the missing mass multiplied by the spectrum ZPA. This method is mathematically rigorous and is considered the only acceptable method to account for high frequency modal contributions ($f \geq f_{ZPA}$) and mass apportioned to system support points.

Kennedy (Reference 3) documented this method and recommended that it be included in Regulatory Guidance. The 1989 revision to the SRP Section 3.7.2, "Seismic System Analysis," (Reference 2) incorporated Kennedy's recommendation as Appendix A. The mathematical details are presented in both References 2 and 3, and are not repeated here. However, the guideline provided in References 2 and 3, that the missing mass contribution needs to be considered only if the fraction of missing mass at any degree of freedom exceeds 0.1, should be eliminated. This guideline does not consider the total mass which is missing, which in the limit could be 10%. In a static analysis this represents a 10% reduction in the applied load. The missing mass contribution should be calculated in all response spectrum analyses, because its potential effect on support reactions is difficult to judge based on the fraction of missing mass. This calculation has been automated in a number of piping analysis codes and does not represent a significant computational effort.

The missing mass contribution to the response spectrum analysis solution represents response which is completely in-phase with the time varying acceleration input and can be scaled to the instantaneous acceleration to obtain its contribution at any specific point in time. This

characteristic is not important in response spectrum analysis because only peak response is predicted. In this case, the ZPA is used to generate the missing mass loading. However, the importance of the missing mass contribution is not limited to response spectrum analysis only. Mode superposition time history analysis is most accurately and efficiently performed by a procedure similar to that employed in response spectrum analysis (Reference 4). Only modes which vibrate at frequencies below f_{ZPA} need to be included in the transient mode superposition solution. The missing mass contribution, scaled to the instantaneous acceleration, is then algebraically summed with the transient solution at the corresponding time to obtain the total solution. This method is more rigorous and accurate than including additional modes in the transient mode superposition solution. Even if additional modes are included, it is still necessary to calculate the missing mass for the excluded, higher frequency modes and system support points. This was quantitatively demonstrated in a separate numerical study.

Use of the Missing Mass method for calculating the contribution of high frequency modes is recommended in Draft ASCE Standard 4 (Reference 9) for both response spectrum analysis (Eqn. 3200-8) and mode superposition time history analysis (Eqn. 3200-5). In Reference 9, this is referred to as the "residual rigid response due to the missing mass."

2.3.2 Static ZPA Method

The Lindley-Yow Method (Reference 13) defines the acceleration of the in-phase response component of all modes to be the ZPA of the response spectrum. As discussed in Section 2.2.1, the algebraical summation of the in-phase response components for all modes (R_r) is equivalent to the static response for a load equal to the total mass times ZPA. When using the Lindley-Yow method, an alternate approach to including the contribution of high frequency ($f \geq f_{ZPA}$) modes is to calculate R_r directly by the Static ZPA method. This eliminates the need for calculation of the missing mass, since it is automatically included in the static analysis of total mass times ZPA. The out-of-phase response component (R_p) is calculated in accordance with the Lindley-Yow method.

A significant result was obtained during the course of this investigation which led to a supplementary study of differences in mass distribution between a dynamic analysis model and a static analysis model. In the approach defined in Section 2.2.1, the dynamic mass distribution is used in constructing the total in-phase response, R_r . In the Static ZPA method, the static mass distribution is used to develop R_r . The out-of-phase response, R_p , is based on the dynamic model mass distribution in both cases. During numerical studies of a piping model, using a BNL version of SAP V adapted for piping analysis, correlation between Lindley-Yow plus Missing Mass and the Static ZPA approach could not be achieved. Further investigation identified the source of the discrepancy to be different treatments of the piping system mass. In the dynamic analysis, the distributed mass of the pipe is replaced by discrete masses at the nodes of the model. In the static analysis, the mass remains distributed along the pipe elements. When the dynamic mass distribution is used in the Static ZPA method, the discrepancy disappears and excellent correlation is achieved. This issue was further evaluated; guidelines for ensuring that the model refinement is sufficient to accurately represent the distributed mass were developed.

2.4 Complete Solution for Response Spectrum Analysis

For the numerical studies conducted as part of this project, three approaches were defined for constructing the complete response spectrum analysis solution. For simplicity, these have been designated Method 1, Method 2, and Method 3, and are defined below. The coefficients C_{jk} are defined by one of the out-of-phase combination methods (Section 2.1.1 through 2.1.6). In the numerical studies, all six methods were tested in conjunction with Methods 1, 2, and 3.

2.4.1 Method 1

Method 1 represents the common method applied to response spectrum analysis since the 1980's. Amplified modal responses ($f < f_{ZPA}$) are combined by SRSS with a correction for closely spaced modes. The contribution of unamplified modal responses ($f > f_{ZPA}$) is calculated by the missing mass method of Section 2.3.1. These two components are then combined by SRSS to produce the total solution. Mathematically, this is represented by

$$R_p = \left[\sum_{j=1}^n \sum_{k=1}^n C_{jk} R_j R_k \right]^{1/2}$$

$n = \text{no. of modes below } f_{ZPA}$ (Eqn. 2-23)

$$R_r = R_{\text{missing mass}}$$

$$R_t = \sqrt{R_p^2 + R_r^2}$$

2.4.2 Method 2

Method 2 introduces the concept of in-phase and out-of-phase modal response components for the amplified modes ($f < f_{ZPA}$). Mathematically, the complete solution is represented by

$$R_{p_i} = R_i * (1 - \alpha_i^2)^{1/2}$$

$$R_{r_i} = R_i * \alpha_i$$

$$R_p = \left[\sum_{j=1}^n \sum_{k=1}^n C_{jk} R_{p_j} R_{p_k} \right]^{1/2}$$

$n = \text{no. of modes below } f_{ZPA}$ (Eqn. 2-24)

$$R_r = \sum_{i=1}^n R_{r_i} + R_{\text{missing mass}}$$

$$R_t = \sqrt{R_p^2 + R_r^2}$$

The method recommended in Draft ASCE Standard 4 (Reference 9) for obtaining the complete response spectrum analysis solution (Eqns. 3200-17 and 3200-18) is essentially equivalent to Method 2.

Method 2 is equally applicable to both the Lindley-Yow (Section 2.2.1) and the Gupta (Section 2.2.3) methods. Only the definition of α_i changes. For the initial numerical study, the Lindley-Yow method was selected for implementation. For the follow-up numerical study, the Gupta method was implemented.

2.4.3 Method 3

Method 3 is a variation of Method 2, which utilizes the Static ZPA Method of Section 2.3.2 to calculate R_r . Mathematically, the complete solution is represented by

$$R_{p_i} = R_i * (1 - \alpha_i^2)^{1/2}$$

$$R_p = \left[\sum_{j=1}^n \sum_{k=1}^n C_{jk} R_{p_j} R_{p_k} \right]^{1/2}$$

(Eqn. 2-25)

$n = \text{no. of modes below } f_{ZPA}$

$$R_r = R_{\text{static ZPA}}$$

$$R_t = \sqrt{R_p^2 + R_r^2}$$

Method 3 is only compatible with the Lindley-Yow method. α_i must be defined by Equation 2-9.

3 SUMMARY OF RESULTS

The qualitative evaluation of modal response combination methods provided the foundation for the subsequent numerical studies, which quantitatively evaluated the strengths and weaknesses of the combination methods by comparison to time history analysis results. Together, the qualitative and quantitative evaluations provided the basis for technical conclusions and recommendations for revision of regulatory guidance. Complete documentation of this project will be published as a NUREG/CR in Spring 1999.

4 REFERENCES

1. U.S. Nuclear Regulatory Commission, "Combining Modal Responses and Spatial Components in Seismic Response Analysis," Regulatory Guide 1.92, Revision 1, February 1976.
2. U.S. Nuclear Regulatory Commission Standard Review Plan for the Review of Safety Analysis Reports for Nuclear Power Plants, NUREG-0800, Section 3.7.2, "Seismic System Analysis," Revision 2, August 1989.
3. Kennedy, R.P., "Position Paper on Response Combinations," Report No. SMA 12211.02-R2-0, March 1984 (Published in Report of the U.S. Nuclear Regulatory Commission Piping Review Committee: Evaluation of Other Dynamic Loads and Load Combinations, NUREG-1061, Vol. 4, pp. B-43 to B-95, April 1985).
4. Gupta, A.K., "Response Spectrum Method in Seismic Analysis and Design of Structures," CRC Press, Inc., 1993.
5. Bezler, P., Curreri, J.R., Wang, Y.K., and Gupta, A.K., "Alternate Modal Combination Methods in Response Spectrum Analysis," NUREG/CR-5627, BNL, October 1990.
6. Wang, Y.K., Subudhi, M., and Bezler, P., "Investigation of the Conservatism Associated with Different Combinations Between Primary and Secondary Piping Responses," NUREG/CR-3086, BNL, January 1983.
7. Bezler, P., Hartzman, M., and Reich, M., "Piping Benchmark Problems Dynamic Analysis Uniform Support Motion Response Spectrum Method," NUREG/CR-1677, Vol.1, BNL, August 1980.
8. American Society of Civil Engineers, "Seismic Analysis of Safety Related Nuclear Structures and Commentary on Standard for Seismic Analysis of Safety Related Nuclear Structures (4-86), September 1986.
9. American Society of Civil Engineers, "Seismic Analysis of Safety Related Nuclear Structures and Commentary," ASCE 4, Revision, Draft, April 1995.
10. Maison, B.F., Neuss, C.F., and Kassaj, K., "Earthquake Engineering and Structural Dynamics, Vol. II, p. 623-647; "The Comparative Performance of Seismic Response Spectrum Combination Rules in Building Analysis," John Wiley & Sons, 1983.
11. Rosenblueth, E., and Elorduy, J., "Responses of Linear Systems to Certain Transient Disturbances," Proceedings of the Fourth World Conference on Earthquake Engineering, Santiago, Chile, 1969.
12. Der Kiureghian, A. "A Response Spectrum Method for Random Vibrations," University of California at Berkeley, June 1980.
13. Lindley, D.W., and Yow, T.R., "Modal Response Summation for Seismic Qualification," Proceedings of the Second ASCE Conference on Civil Engineering and Nuclear Power, Vol. VI, Paper 8-2, Knoxville, TN, September 1980.
14. Hadjian, A.H., "Seismic Response of Structures by the Response Spectrum Method," Nuclear Engineering and Design, Vol. 66, no. 2, pp. 179-201, August 1981.

Post Test Analysis of a PCCV Model Subjected to Beyond-Design-Basis Earthquake Simulations

R. J. James, Y. R. Rashid
ANATECH Corp.

J. L. Cherry
Sandia National Laboratory

N. Chokshi
USNRC

ABSTRACT

A scaled model Prestressed Concrete Containment Vessel (PCCV) was tested up to ultimate failure under simulated earthquake loadings at the Nuclear Power Engineering Corporation's (NUPEC) Tadotsu Engineering Laboratory in Japan. The mixed-scale model was first subjected to a series of design-level earthquakes, and then the magnitudes of the earthquakes were increased in several stages until the cylinder walls catastrophically failed in shear. Under sponsorship from the U.S. Nuclear Regulatory Commission (USNRC), state-of-the-art, nonlinear dynamic finite element analyses were completed for the PCCV structure. Both pre- and post-test analyses were performed. This paper summarizes post-test analyses that correspond to the larger-than-design-basis seismic tests. Under these "severe" conditions, the concrete cracks, spalls, and crushes while the reinforcing bars and liner under go extensive plastic deformation during the repeated dynamic load cycling. The analytical results are compared to the measured structural response of the scaled-model structure. The USNRC's principal objective of this work has been to evaluate how well state-of-the-art analyses can predict the structural response and eventual failure of a prestressed concrete structure under seismic loadings that are considerably larger than the structure was designed for.

1. INTRODUCTION

The seismic behavior of a prestressed concrete containment vessel (PCCV) is the object of experimental and analytical investigations in a collaborative program between the United States Nuclear Regulatory Commission (NRC) and the Nuclear Power Engineering Corporation (NUPEC) of Japan.

NUPEC's primary objective is to demonstrate the capability of the PCCV to withstand the design basis earthquake with significant safety margins against major damage or failure. The test structure is a mixed-scale PCCV model, subjected to seismic simulation tests using the high performance shaking table at the Tadotsu Engineering Laboratory [1]. Acceleration time histories were developed to be representative of typical design level earthquakes, and then scaled to excite a response in the model that will be similar to that in the actual structure. The test program included design-basis and failure-level earthquakes; one of the design basis tests also included internal pressurization to simulate a loss-of-coolant-accident (LOCA). The margin of safety was determined by subjecting the model to progressively larger seismic motions until structural failure occurred. The fundamental frequencies of the test model, which give a measure of the damage sustained by the model, were determined after each test.

NRC's objective is to evaluate the capabilities of the state-of-the-art concrete structural analysis methods for predicting the dynamic behavior of concrete containments subjected to design-level and failure-level earthquakes, and to identify areas of needed improvements. A large amount of structural response data was obtained for the PCCV scale model subjected to design-basis, as well as beyond-design-basis, earthquakes. This included earthquake motions, which excited a structural response, in the linear range and progressively stronger motions where significant structural damage began to accumulate up to major structural damage and final failure. Test data was obtained for horizontal-only earthquake input, vertical-only input, simultaneous horizontal and vertical, and simultaneous horizontal and vertical combined with design pressure. Post-test finite element analyses were performed that corresponded to several of these tests, and the calculated responses were very similar to the measured results.

The analytical program included two series of analyses: pre-test predictions and post-test verification analyses. The pre-test analyses used target input acceleration time histories and were reported on in the 1997 Water Reactor Safety Meeting [2]. The post-test calculations used measured acceleration time histories and are described in this paper.

2. PCCV TEST MODEL

The PCCV test model, as developed by NUPEC, used mixed scaling for practical considerations: overall geometry was scaled at 1:10, while the concrete wall thickness was scaled at 1:8; the steel liner and "T" anchorage system were scaled at 1:4; and the dome portion of an actual PCCV structure was replaced with a thick flat concrete cap. Weights were attached to the top slab to match the response of the model to that of a prototype structure for shear stress in the wall at the wall-basemat juncture. The test model has a cylindrical barrel with an ID of 4.3 m, a wall thickness of .163 m, and a height of 3.43 m. It is cast on a 9-m square by 1-m thick basemat that is rigidly and securely bolted to the shake table. The top cap is 1-m thick with weights bolted on the top and bottom surfaces and around the outer edge. The cylindrical portion has a 0.600-m diameter equipment hatch (EH) penetration, four main steam lines (MS) and four feedwater (FW) penetrations, and an airlock (AL) penetration. The hoop tendons are anchored at two longitudinal buttresses that are 180° apart. The direction of horizontal shaking is along the diametric line intersecting the two buttresses. The basemat and supporting frames weigh about 260 metric tons, the cylindrical portion weighs 63 metric tons, and the upper section with the added mass weighs 474 metric tons.

3. COMPUTATION AND MATERIAL MODELS

The calculations are performed with the ANACAP-U concrete material model [3] coupled to the ABAQUS general purpose finite element program [4]. In order to simulate damage accumulation in the test series, the nonlinear dynamic analyses were performed in the same sequence as the tests. Figures 1 through 6 illustrate the finite element model used for these calculations. Only a ring section of the basemat around the wall junction was modeled, as shown in Figure 1, with input acceleration time-histories prescribed for all the nodes on the external surfaces of the basemat ring. The portions of the basemat that are rigidly bolted to the shake table are not modeled. The basemat ring is modeled to approximate the area from the wall junction to the first set of bolts that secure the basemat.

In order to reduce the size of the grid, the geometry of the test specimen was assumed to be symmetric about a vertical plane through the two buttresses, and the half containing the equipment hatch was used. The hatch is a large penetration with a thickened section of the PCCV wall and has a thicker liner and additional reinforcement. The concrete finite element grid was developed to provide computational economy while adequately capturing the critical response of interest. The final model, which is shown in Figure 1, has 12,000 degrees of freedom. The PCCV liner, shown in Figure 2, is modeled as membrane (plane-stress) elements fully bonded to the concrete, and are thus strain-compatible with the concrete. All tendons and reinforcement are explicitly modeled, as shown in Figures 3 through 6. The hoop tendons are prestressed to apply about 1138 psi (80 kg/cm²) compression to the concrete, and the axial tendons are prestressed to apply about 1067 psi (75 kg/cm²) compression to the concrete including gravity.

Plasticity relations for the liner, reinforcing bars, and prestressing tendons are included in the material model. The model assumes that the top section and all attached masses of the test specimen remain elastic, and no cracking or compressive yielding of the concrete in this top section is considered in the calculation. The attached masses are modeled with lead material encased in steel shells (as constructed) to capture the distribution of inertial loads and any rocking that may develop. Figure 2 shows the steel cases for the attached weights and the steel plates embedded in the top section.

The material model used for the concrete in the PCCV wall and basemat is the ANATECH concrete material model, known as ANACAP-U [3]. The material properties used for the concrete are as follows:

| | |
|----------------------|---|
| Modulus | 3.4E6 psi (2.39E5 kg/cm ²) |
| Poisson's Ratio | 0.19 |
| Compressive Strength | 5633 psi (396 kg/cm ²) |
| Fracture Strain | 158.7E-6 |
| Weight Density | 150 lb/ft ³ (0.0024 kg/cm ³) |

Based on Raphael's formula [5], $f'_c = 1.7 f'_c{}^{2/3}$ (in units of psi), the material is assumed to have a tensile strength of 540 psi (38 kg/cm²). Under uniaxial compression, the model assumes that the material will reach its compressive strength at 3000E-6 strain and begins to soften under additional load. Figure 7 illustrates the uniaxial compressive stress-strain relationship and the shear retention model used for the PCCV concrete and shows the hysteretic and shear degradation behavior that is typical of concrete under high cyclic compression and shear loading.

4. POST-TEST ANALYSIS RESULTS

Comparison of the actual input acceleration records with the target accelerations used for the pre-test analyses revealed major differences between the two inputs for the beyond-design-basis and failure-level motions. Table 1 shows the maximum for the two sets of input accelerations. Furthermore, the basemat acceleration records, which form the input to the analysis, show significant rocking, as shown in the table. This vertical motion was totally missing from the target accelerations used in the pre-test analysis. As already mentioned, the post-test analyses described in this paper use the measured input accelerations, and space does not permit comparison with the pre-tests analyses. For the same reason, detailed description of the post-test analyses cannot be given, and only representative results are provided.

Table 1. Target vs. Actual Basemat Acceleration for the Analyzed Tests

| Post-Tests Analyzed | Horizontal Acceleration (g) | | Vertical Acceleration (g) | | |
|---------------------|-----------------------------|----------|---------------------------|---------|---------|
| | Target | Measured | Target | Uniform | Rocking |
| S1(H+V) | 0.28 | 0.30 | 0.13 | 0.18 | 0.20 |
| S2(H+V) | 0.43 | 0.44 | 0.21 | 0.30 | 0.30 |
| S2(H+V)+LOCA | 0.28 | 0.29 | 0.13 | 0.22 | 0.22 |
| 2.0S2(H) | 0.86 | 1.46 | - | 1.17 | 1.51 |
| 3.0S2(H) | 1.29 | 2.72 | - | 2.10 | 2.57 |
| 3.3S2(H) low freq. | 1.42 | 2.56 | - | 1.33 | 1.78 |
| 5.0S2(H) | 2.15 | 3.53 | - | 1.82 | 2.37 |

4.1 Design Level Analyses

The sequence of tests chosen for the post-test calculations is S1(H+V), S2(H+V), and S1(H+V)+LOCA, where H and V stand for horizontal and vertical, respectively. The internal pressure used in the analysis for the loss of coolant accident (LOCA) is 56.9 psi (4kg/cm²). The loading for the analysis sequence was determined from the recorded accelerometer data from the top surface of the basemat. Figure 8 illustrates the input accelerations for the S1(H+V)+LOCA, and Figure 9 illustrates the cracking patterns for the same event. The calculated crack widths are of the order of 0.14 mm. The calculation predicts vertical

cracks near a buttress with small crack widths of about .01 mm. This type of cracking is also observed in the test. The cracking observed in the test near the top section does not develop in the calculations.

The global response of the PCCV model is illustrated in Figures 10 and 11, which show the horizontal displacement and acceleration of the top mass respectively. It is apparent that, for these design level calculations, the analytical simulation is under-predicting the response. This is attributed to the damping used in the analyses. After these calculations were performed, the test data was analyzed by NUPEC to determine the level of damping exhibited in the test. This evaluation indicated the overall damping to be about 1% initially and increased to about 1.5% during the design level tests. It continues to increase to about 3.5% as further cracking developed in the failure level series of tests. Thus, the 3% uniform damping used in the calculation was too high for the design level tests, which contributed to the under-prediction of the response.

4.2 Failure Level Analyses

For the PCCV failure level testing, the test plan called for determining the seismic margin by subjecting the model to increasing multiplies on the S2(H) level magnitudes until structural failure occurred. For these tests, only the horizontal input motion was planned. However, increasing the horizontal amplitude caused substantial vertical feedback and rocking of the PCCV test model, resulting in rather substantial vertical acceleration input on the basemat, as summarized in Table 1.

Good agreements were obtained in general, but the best agreement was obtained for the 3.3S2(H) test. The results for this test are illustrated in Figures 12 through 15, which depict the horizontal and vertical displacements and accelerations for the top mass. The excellent agreement obtained for this test is attributed in part to the fact that the 3% damping used in the analysis is consistent with the level of damage in the structure.

The PCCV failed during the 5.0S2(H) test, as can be seen from Figures 16 and 17. The test model began to fail around 4.9 seconds at the edge of the equipment hatch at cylinder mid-height, followed immediately by failure in the wall near the top section at the wall thickness transition. A band of the concrete at the top rubblized and fell out over the next 0.8 seconds, resulting in settlement of the top section with buckling and tearing of the liner. There is also some spalling damage around the buttresses at the basemat juncture. Figure 18 illustrates the areas of damage in the test model after failure during the 5.0S2(H) test. The calculated response for this test simulation agrees reasonably well up to the point of failure. Beyond that, the physical loss of material prevented further correlation.

4.3 Failure Criterion

By examining the state of the solution around the time of failure, the conditions leading to shear failure can be identified in the calculations. From this evaluation it is proposed that shear strains of 0.5% or higher acting over at least 80% of any cross section is a state of impending structural failure. Figure 19 illustrates the shear strain contours at the time of failure in the test with the minimum range set at 0.5%. The large band of shear strain greater than or equal to 0.5% at mid-height extending across the section is evident. This hypothesis was checked and confirmed with similar calculations for the shear wall tests in Ref. 6, and Figure 20 shows the calculated shear strains in the NUPEC shear wall at the time of failure. The calculations would predict failure at the correct time and at approximately the correct locations in the shear wall test when the calculated shear strains reached at least 0.5% across a section of the wall. This failure measure with an appropriate uncertainty band is proposed as a general failure criterion for shear structures subjected to severe seismic events.

4.4 Summary of Results

Table 2 summarizes the fundamental frequency shift in the post-test calculations as damage accumulates. Table 3 summarizes the peak horizontal accelerations and displacements of the top mass in the post-test calculations relative to the test data. Figure 21 shows the comparison of calculated peak horizontal response and test data graphically. For the design level tests, the post-test calculations under-predict the response due to the high value of damping (3%) used in the analysis throughout. An evaluation of test data indicates that damping in the test specimen varied from 1% initially, increasing to 1.5% during the design level tests, to 3.5% for the failure level tests. The agreement in the failure level post-test calculations is quite good compared to the test data. The overall behavior and peak response agree well with the test.

Table 2. Summary of Fundamental Frequency Shift in Post-Test Models

| Post-Test Models | Test Data | | Post-Test Model | |
|--------------------|-----------|--------------------------|-----------------|--------------------------|
| | Hz | % Reduction from Initial | Hz | % Reduction from Initial |
| S1(H+V) | 10.8 | - | 11.3 | - |
| S2(H+V) | 10.3 | 4.6 | 11.3 | 0.0 |
| S2(H+V)+LOCA | 9.9 | 8.3 | 11.3 | 0.0 |
| 2.0S2(H) | 9.0 | 16.7 | - | - |
| 3.0S2(H) | 8.8 | 18.5 | 9.1 | 19.5 |
| 3.3S2(H) low freq. | 8.4 | 22.2 | 8.8 | 22.1 |
| 5.0S2(H) | -7.0 | 35.2 | 7.0 | 38.1 |

Currently, no reliable shear failure criterion exists for seismic-resistant structures. Such a shear failure criterion was developed in the course of this investigation, as described earlier, by examining the analytical and experimental results.

Table 3. Summary of Post-Test Horizontal Response of Top Mass

| Post-Test Models | Test Data | | Post-Test Model | |
|--------------------|-----------|------------|-----------------|------------|
| | Acc. (g) | Disp. (mm) | Acc. (g) | Disp. (mm) |
| S1(H+V) | 1.19 | 2.78 | 1.02 | 2.16 |
| S2(H+V) | 1.57 | 3.94 | 1.31 | 2.87 |
| S2(H+V)+LOCA | 1.03 | 2.73 | 0.82 | 1.80 |
| 2.0S2(H) | 2.70 | 12.62 | - | - |
| 3.0S2(H) | 3.37 | 19.60 | 3.67 | 18.00 |
| 3.3S2(H) low freq. | 3.07 | 18.32 | 3.24 | 14.90 |
| 5.0S2(H)* | 3.34 | 25.56 | 3.57 | 24.00 |

*NUPEC reports peak acceleration of 3120 gal and peak displacement of 26 mm. The numbers reported here are taken from response records.

5. CONCLUSIONS AND LESSONS LEARNED

The response of the structure in the range of design level earthquakes was not sensitive to the small differences between the target input accelerations and the actual as-measured accelerations; both pre-test and post-test analyses gave comparable results that agreed reasonably well with the measured data. However, for larger input motions, particularly horizontal-only motions where the vertical motion of the basemat was not controlled, greater differences occurred between the pre-test and post-test results. In fact, it was necessary to perform the post-test analyses for horizontal-only tests, with both input accelerations, horizontal and vertical, as measured by the accelerometers mounted on the basemat. This is because a stronger-than-expected feedback vertical component occurred in the horizontal-only motion, which was not subject to control in the test. Some analytical difficulties in processing the basemat acceleration data for input to the analysis model needed modeling resolutions.

The concrete modeling software used to perform the analysis contained two options for shear resistance: a nominal resistance model based on material-property laboratory tests, and a reduced-resistance model. The latter model is intended for structures with pre-existing damage, a condition that was judged to exist in the test because of the progressively introduced damage in each test, and not every test was analyzed. Both models, however, under-estimated the shear resistance of the structure at the higher load levels, with the latter (reduced-resistance) model showing significant under-prediction. This led to one of the more important findings of the test program, namely, that prestressed concrete containments are robust in

resisting shear even with significant prior damage. Using these findings, a modification of the standard shear model was introduced by adjusting the shear modulus dependence on the crack-opening strain to provide higher shear resistance for narrow cracks. This is the only modification that was introduced in the concrete material model for the post-test analyses. This modification affects the higher-level seismic motions more than lower-level motions where differences between target and actual accelerations had the greater effect on the results.

The second significant finding from this analytical program is the effect of damping on the response. The post-test analyses used a uniform damping value of 3%; however, a close examination of the data and analysis results indicated that the damping, which is a manifestation of damage, is time dependent. The a-priori selection of a single value for an event, which is the current practice, can either underestimate or over-estimate the response. This effect was illustrated by repeating the analysis for one of the time histories in the test series, where a damping value ranging from 1% to 5% is applied locally at the cracked integration points by invoking a crack consistent damping model which represents damping as function of the crack status. The resulting agreement between the analysis and the test was excellent. Unfortunately, however, time and budget constraints did not allow the repetition of the calculations using this time-dependent and damage-dependent representation of damping.

Finally, the testing program provided an opportunity to develop a shear-failure measure for concrete structures subjected to severe seismic motions. Analytical interpretations of the test results indicate that impending shear failure of the structure would occur at a shear strain value of 0.5% (with a suitable uncertainty band) averaged over the entire cross-section of the structure. This value was also confirmed by an analysis of a seismic test of a shear wall that was previously conducted by NUPEC at Tadotsu. This is proposed as a preliminary shear failure criterion pending further verification using additional tests and analyses. This being a structural measure rather than a material property measure, it can only be applied through the post-processing of the analysis results. Further work is needed to adapt the 0.5% structural measure to a concrete material model criteria that would analytically trigger the structural failure.

In summary, the agreement between the calculated time histories and the measured data records is generally good, with instances of poor agreement for some of the gauges and excellent agreement for others. Much better agreement occurs for global measures of response rather than response that is directly affected by local concrete conditions. Very good agreement was obtained for the test with three times the design basis earthquake. This gave an indication that the 3% damping was perhaps a good

value for the damage level in that test. The force-displacement curve for the test series which is a plot of the peak resistance force (mass times acceleration) vs. the maximum horizontal displacement for the tests show very good agreement between analysis and test throughout the test series, with the analytical curve lying consistently above the experimental curve. The higher resistance behavior predicted by the analysis is in part due to the fact that not all of the tests were analyzed, which resulted in lower damage accumulation in the analysis than actually experienced by the structure. The level of agreement between test and analysis achieved in this program is sufficient indication that existing analysis capabilities, with the level of modeling sophistication used in the present analysis, can be relied upon to predict the behavior of concrete containment structures.

6. ACKNOWLEDGMENTS

The analysis work was funded by the USNRC through Sandia National Laboratories Contract AS-9001. SNL is operated by the United States Department of Energy under Contract DE-AC04-94AL 85000. The results described herein are based on analytical predictions performed at ANATECH and do not necessarily reflect the opinion of the USNRC or NUPEC.

7. REFERENCES

1. Y. Sasaki, S. Tsurumaki, H. Akiyama, K. Sato, H. Eto, "Seismic Proving Test of a Prestressed Concrete Containment Vessel," ASME/JSME Joint PVP Conference, San Diego, California, July 1998.
2. Y. R. Rashid, et al., "Seismic Tests of a Prestressed Concrete Containment Vessel; Part 2 - Analytical Investigations," Transactions of the 25th Water Reactor Safety Informational Meeting, Bethesda, Maryland, October 1997.
3. R. J. James, et al., *ANACAP-U, ANATECH Concrete Analysis Package, Version 2.5, User's Manual*, ANATECH Corp., San Diego, September 1997.
4. D. Hibbitt, et al., *ABAQUS/Standard, Version 5.6, User's Manual*, Hibbitt, Karlsson & Sorensen, Inc., Pawtucket, Rhode Island, 1997.
5. J. M. Raphael, "Tensile Strength of Concrete," *ACI Journal*, 82-17, March-April 1984.
6. Y. J. Park, C. H. Hatmayer, "Finite Element Analyses for Seismic Shear Wall International Standard Problem," U.S. NUREG/CR-6554, Nuclear Regulatory Commission, April 1998.

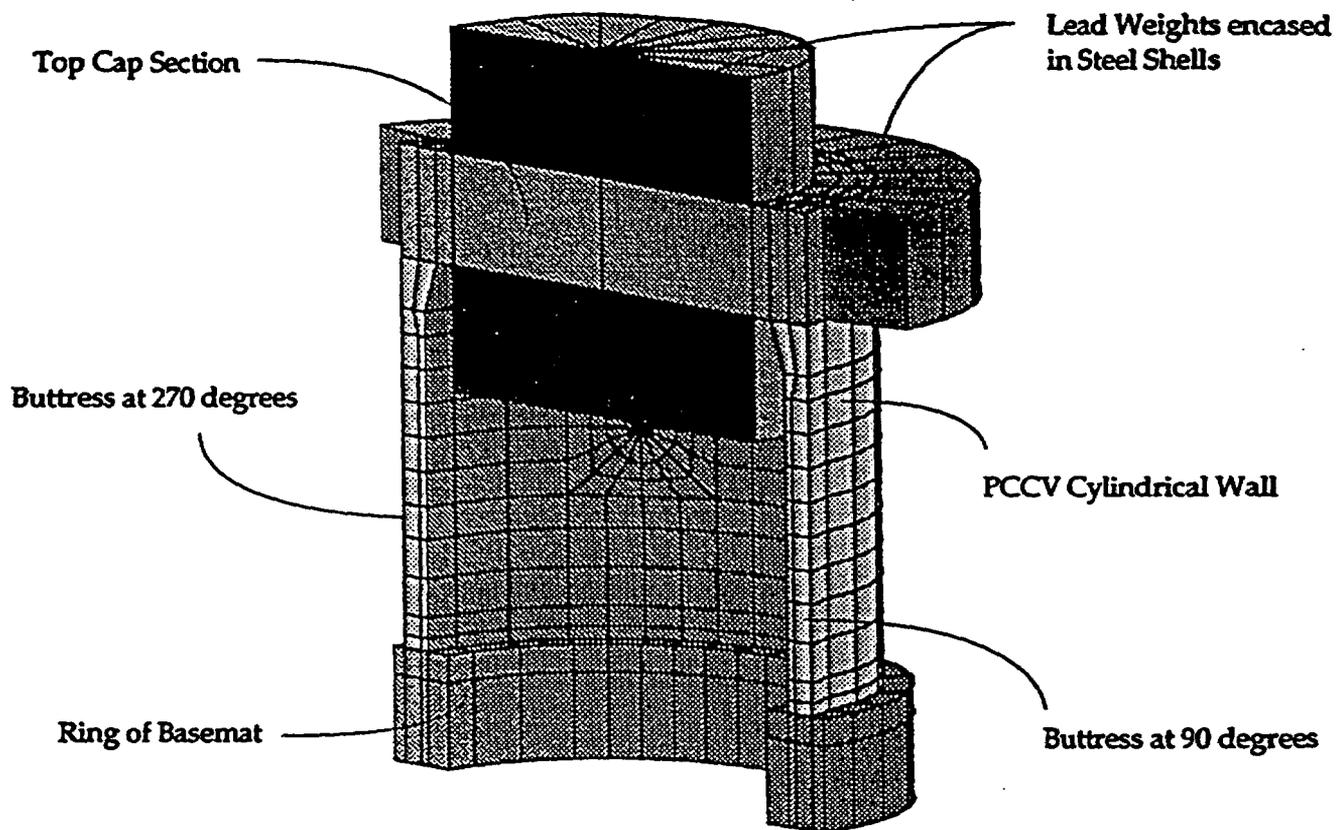


Figure 1. 3D Finite Element Model for Seismic Calculations

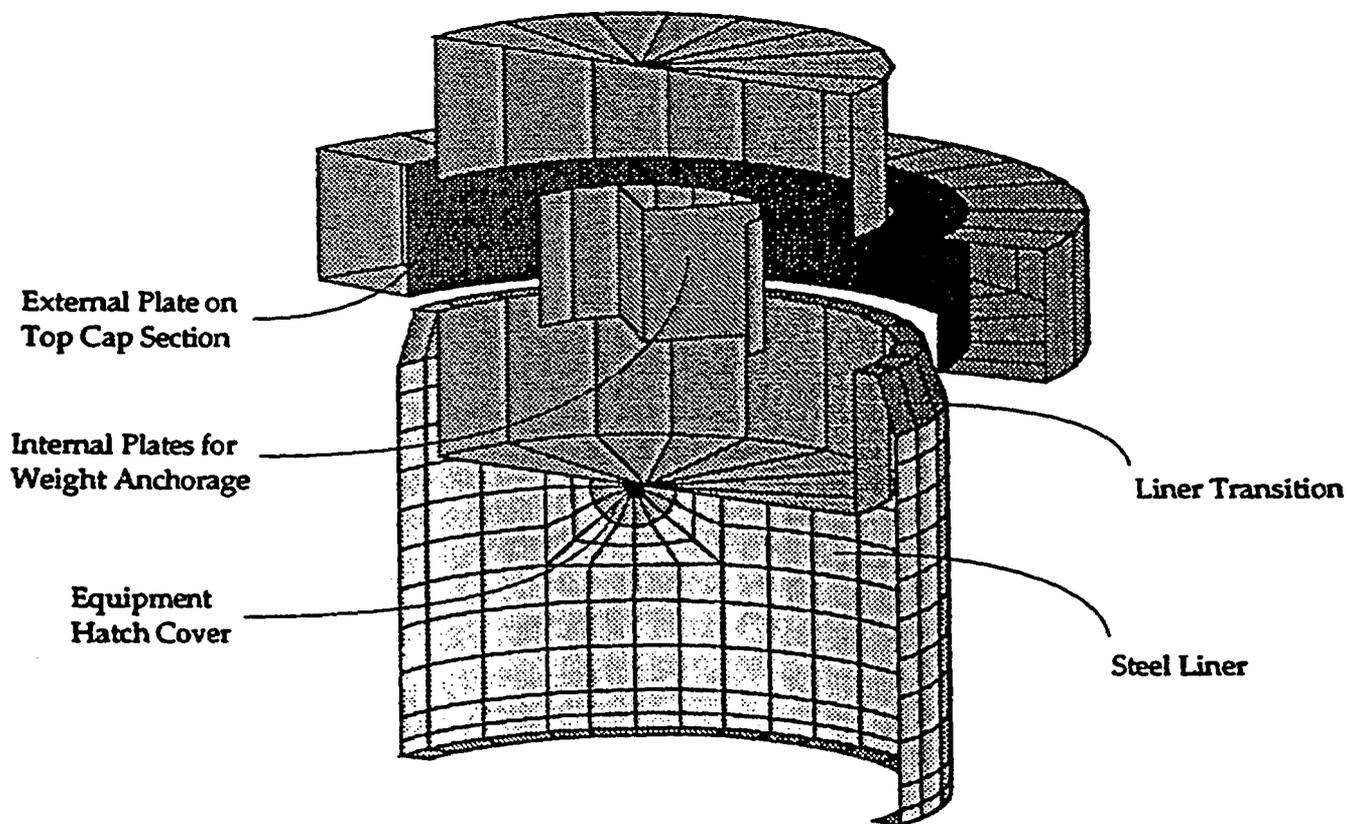


Figure 2. Modeling for the Liner and Steel Plates

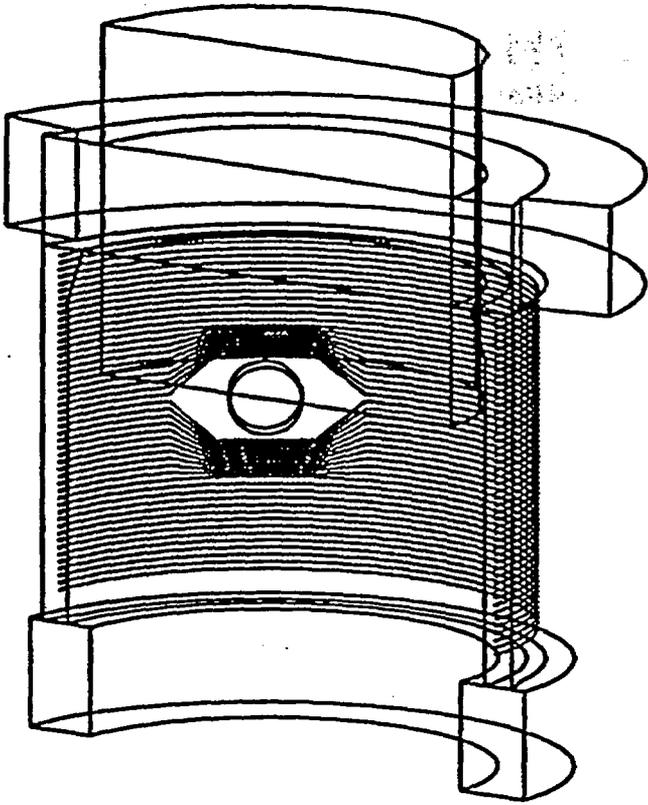


Figure 3. Modeling of Hoop Tendons

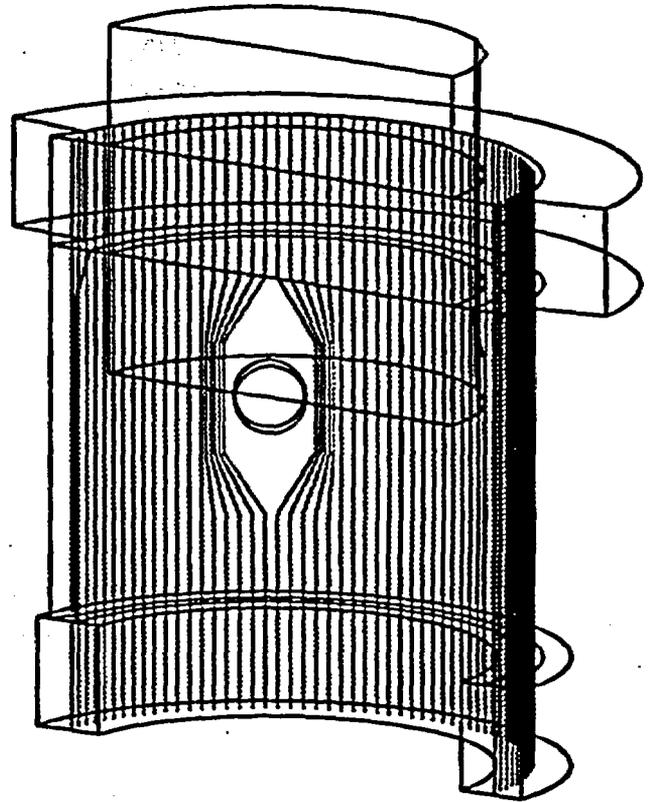


Figure 4. Modeling of Axial Tendons

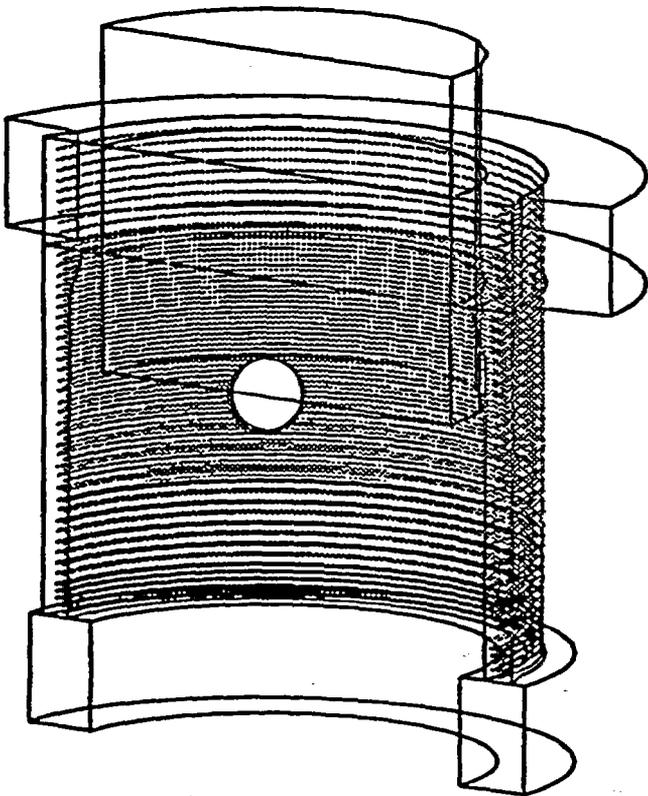


Figure 5. Modeling of Hoop Rebar

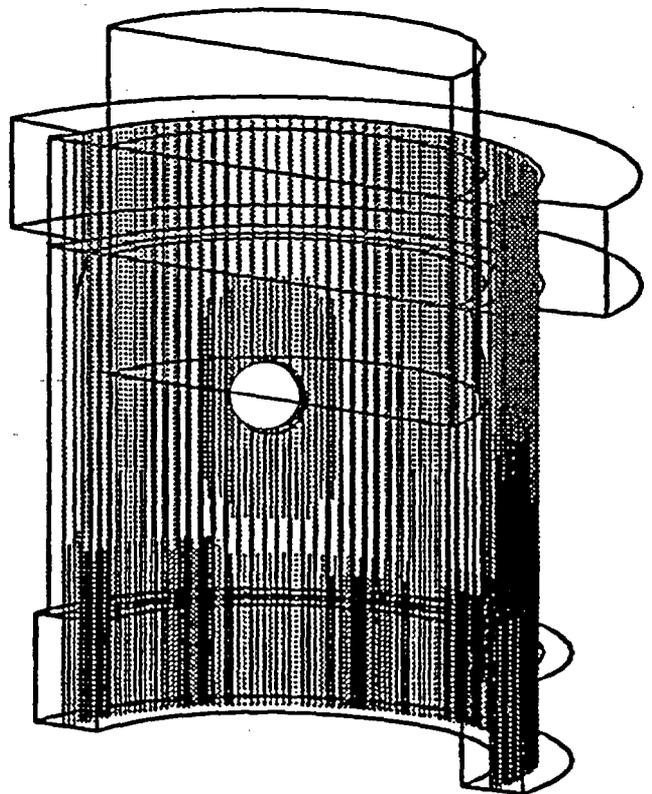


Figure 6. Modeling of Axial Rebar

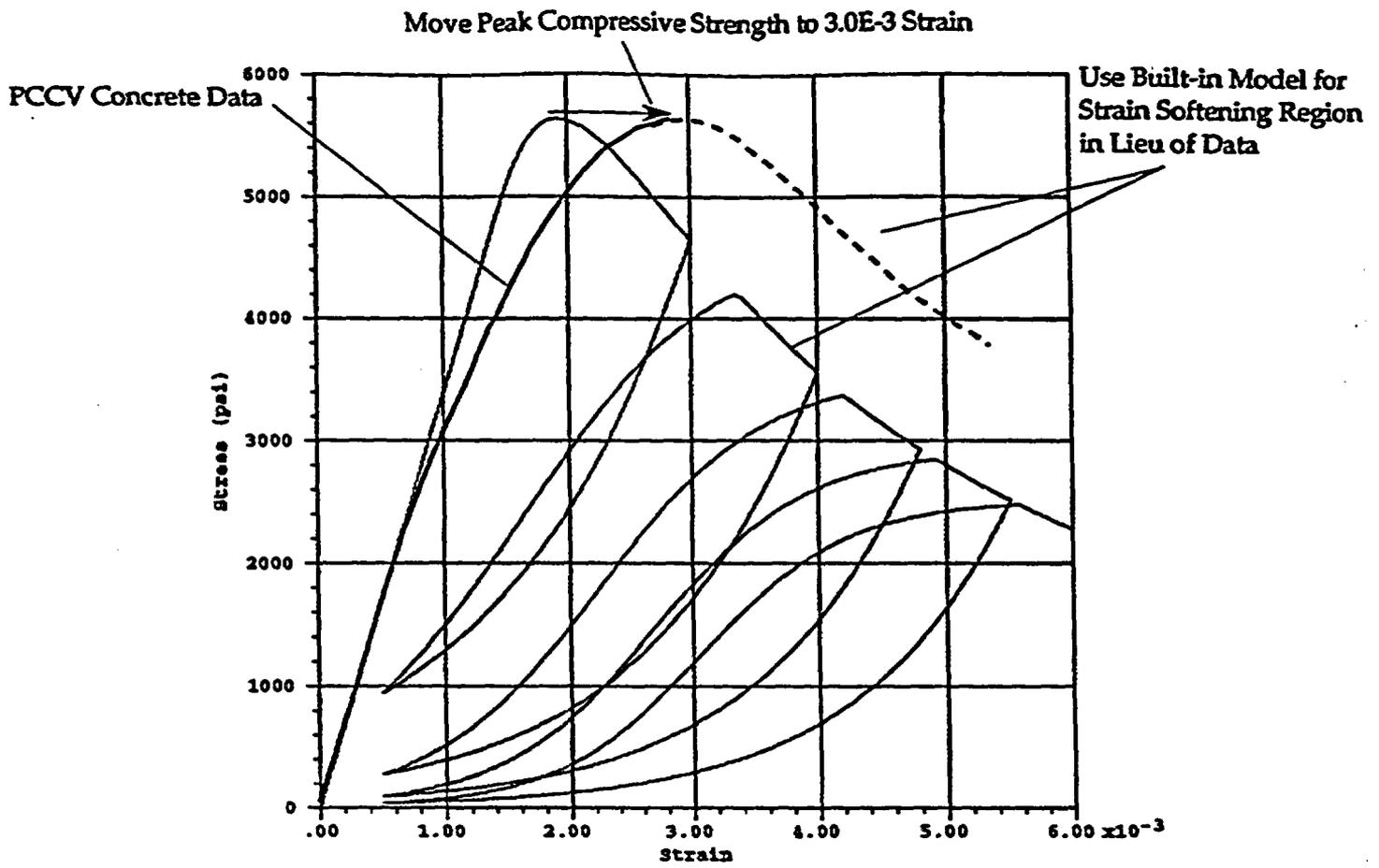


Figure 7a. Compressive Strength Model Adjusted for PCCV Concrete

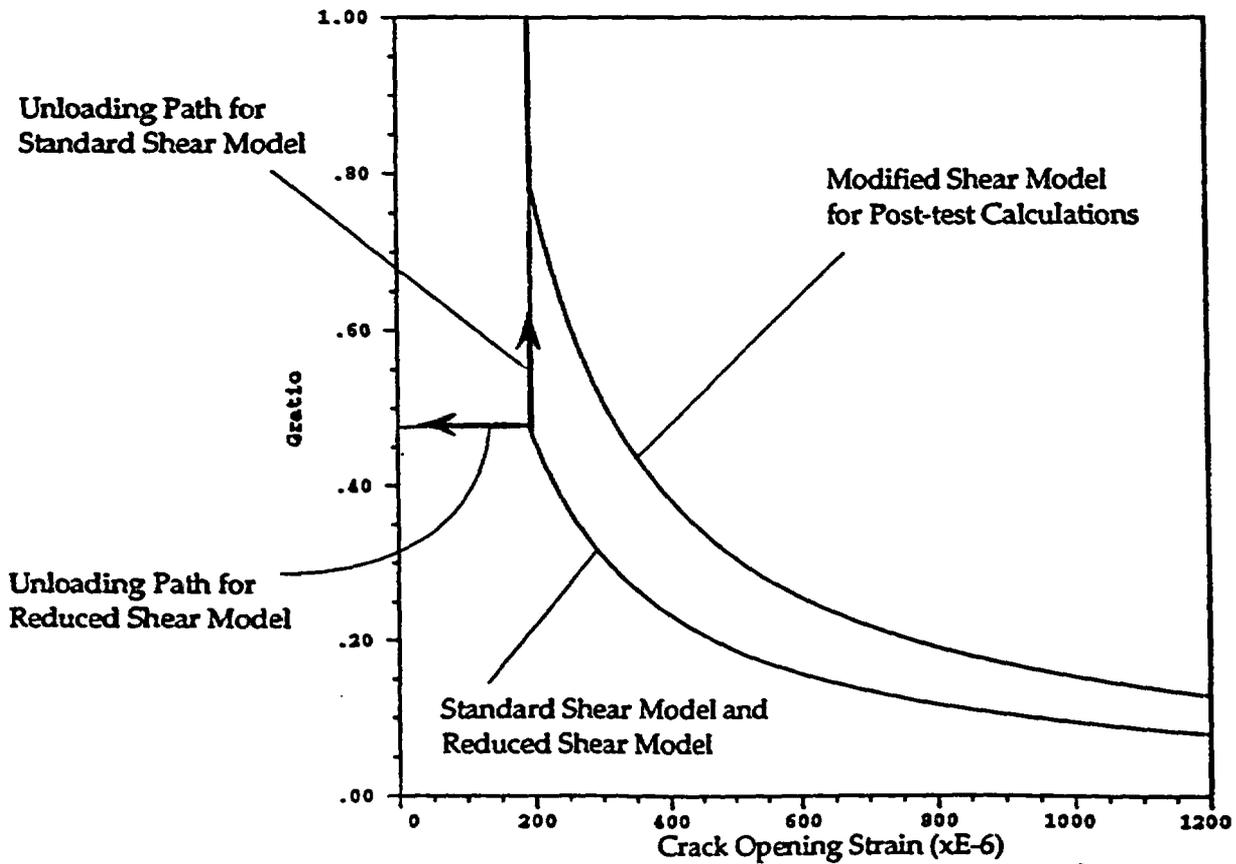


Figure 7b₉₂ Shear Retention Modeling

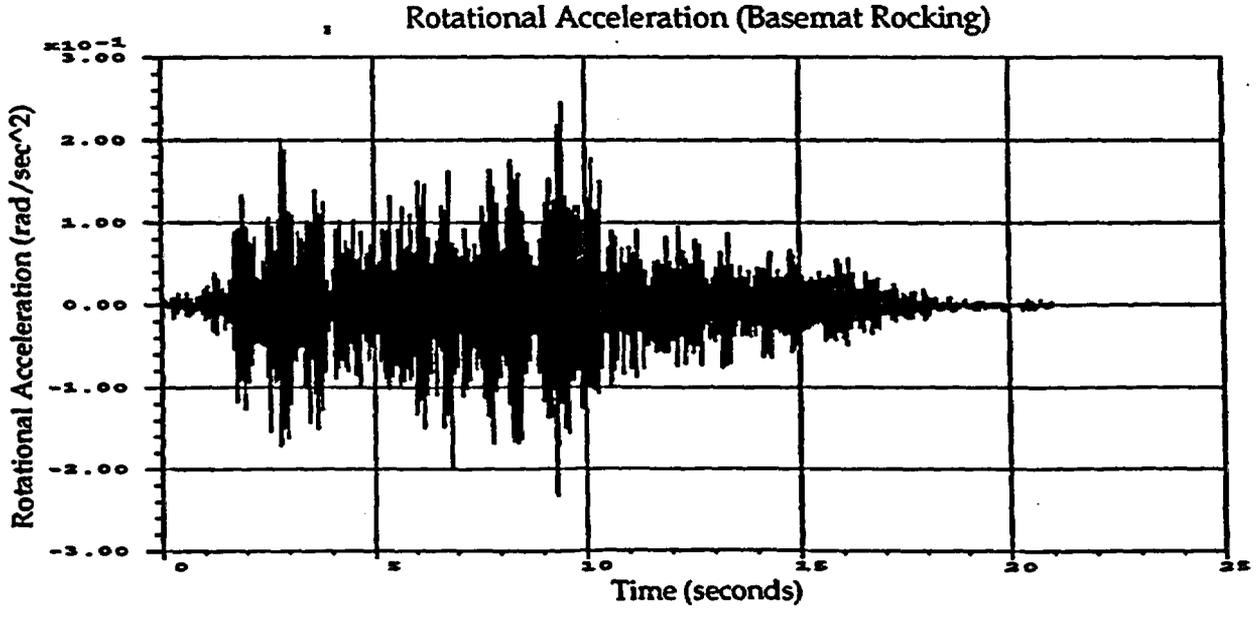
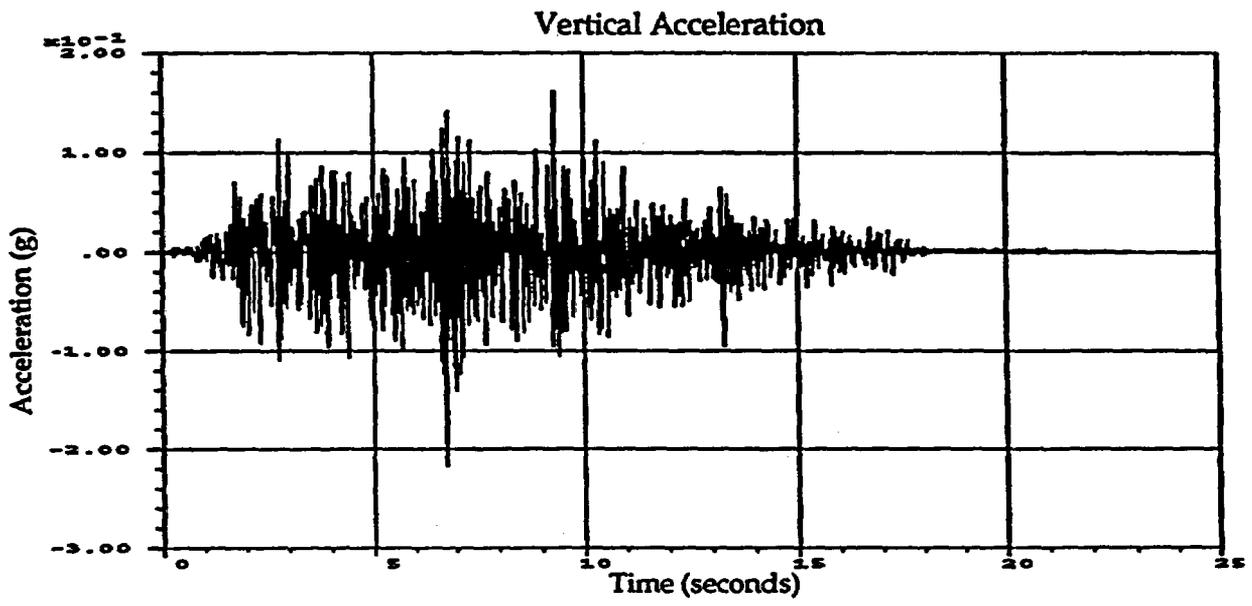
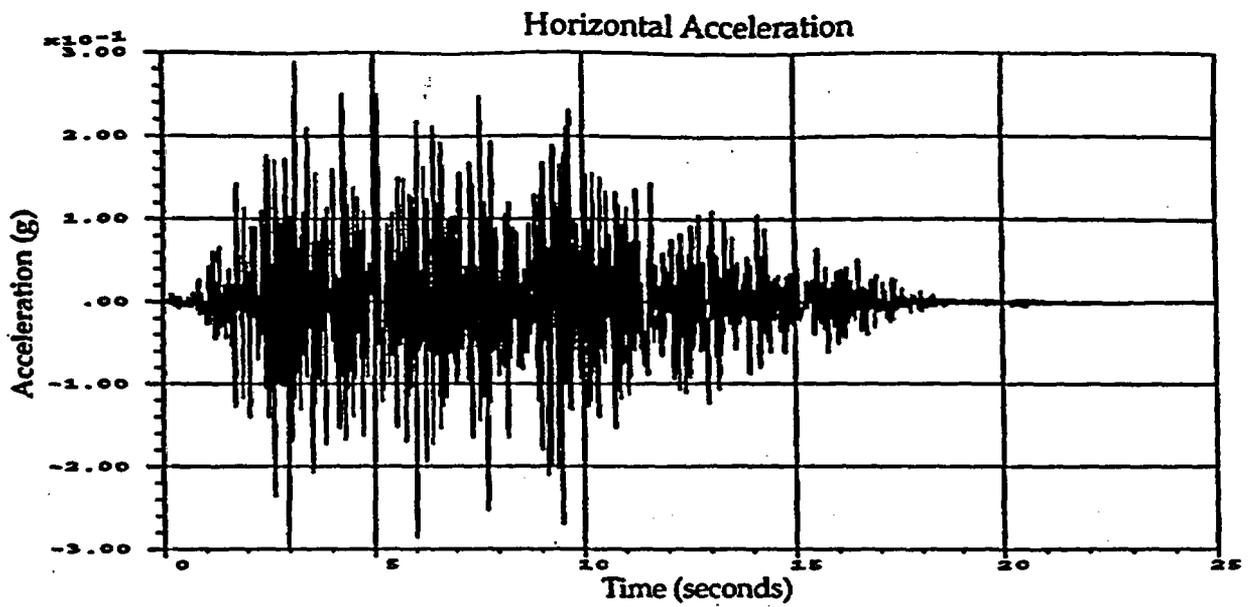


Figure 8. Post-Test Input Acceleration Records for S1(H+V)+LOCA
93

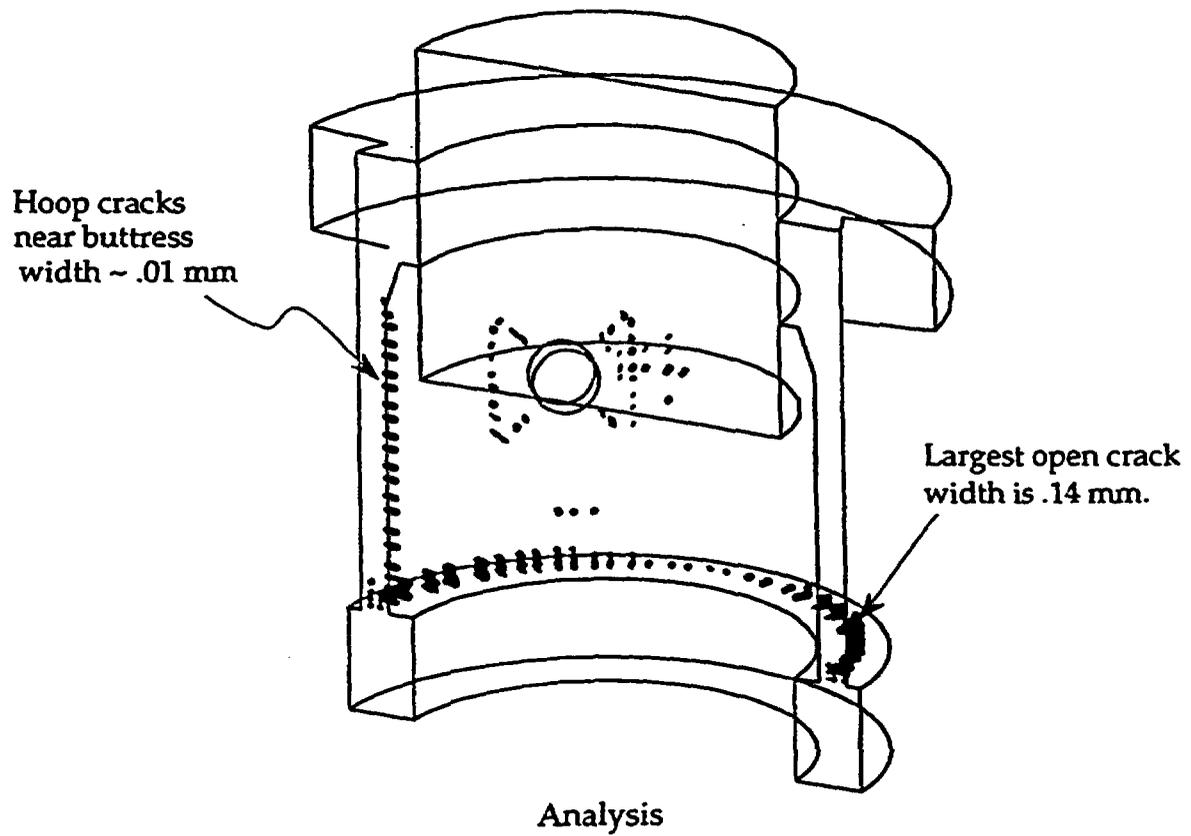
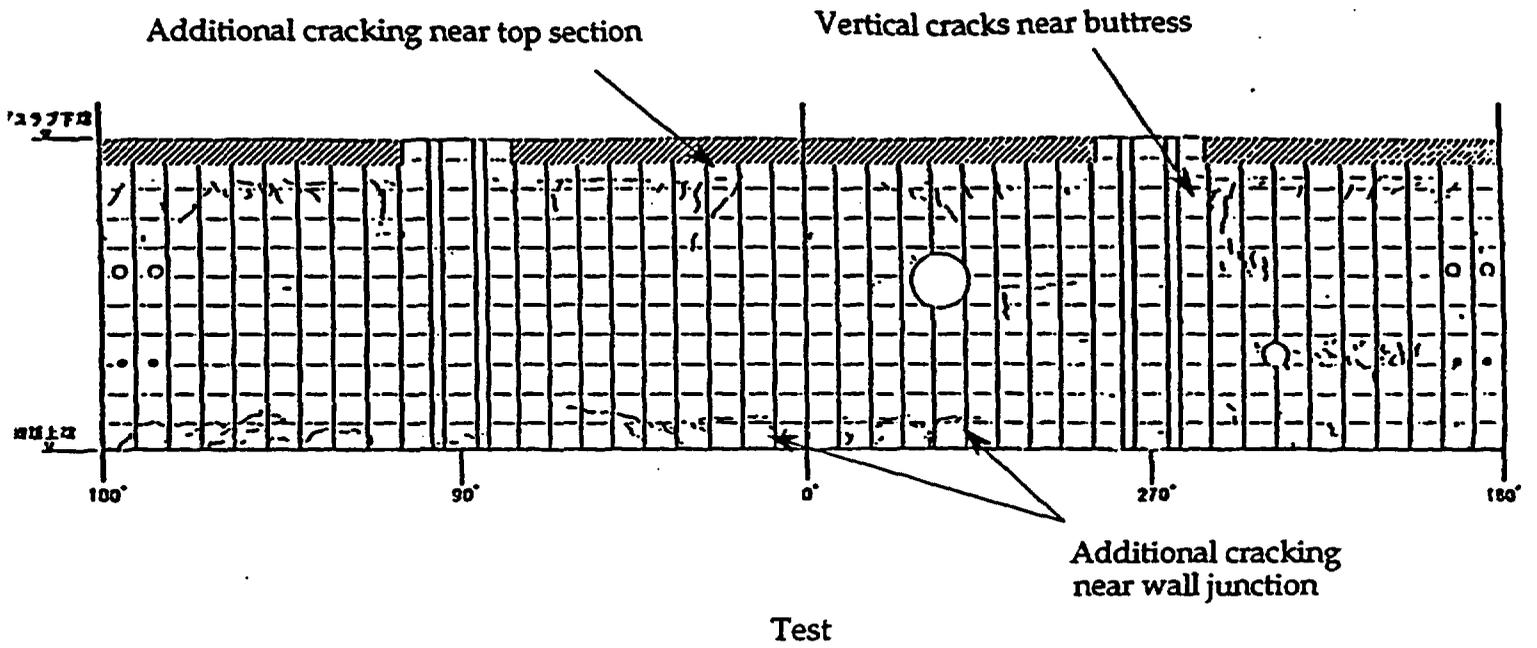
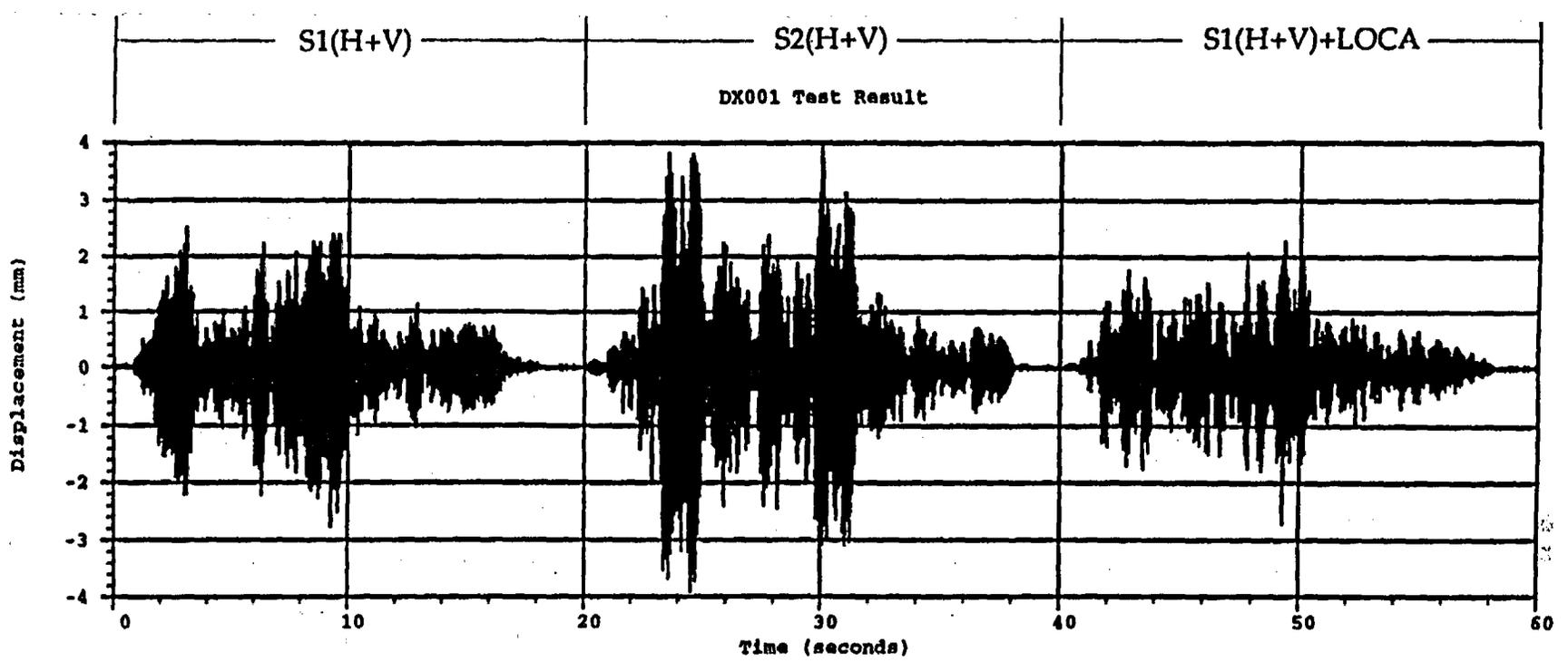


Figure 9. Cracking Patterns After S1(H+V)+LOCA, Post-Test Compared to Data



Post Test Analysis Result

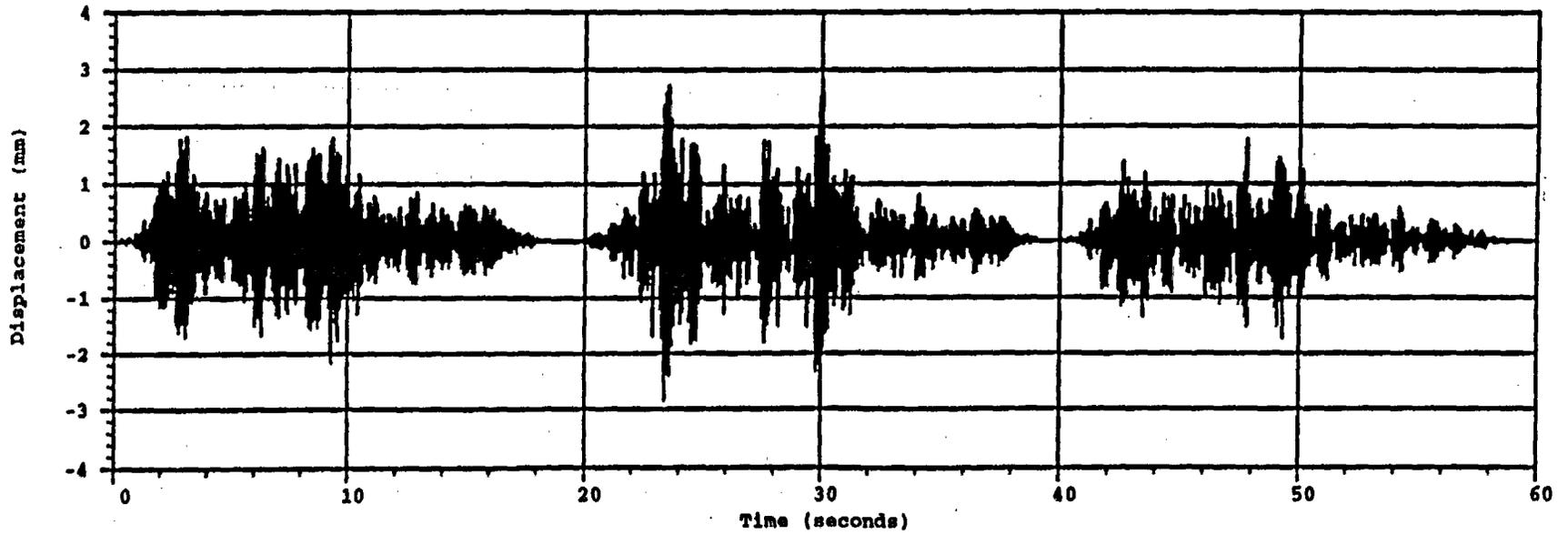
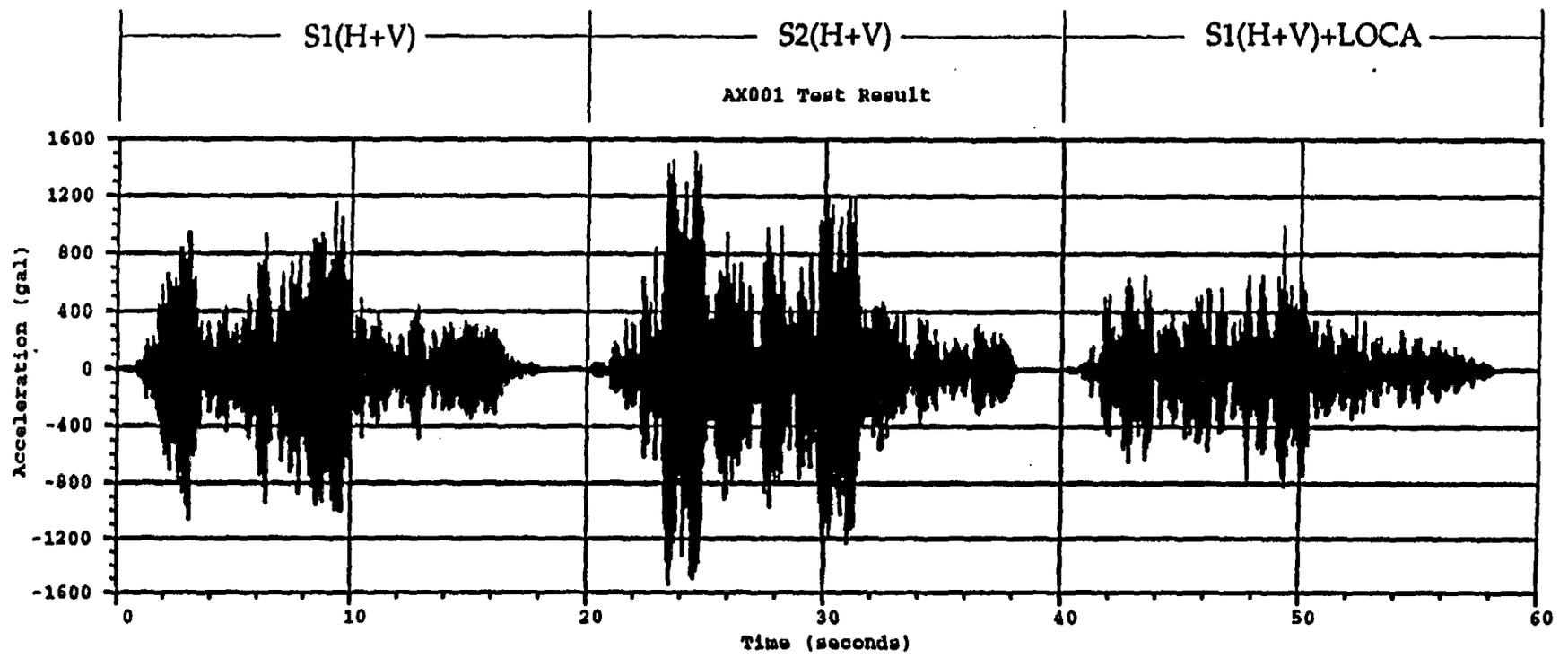


Figure 10. Relative Horizontal Displacement of Top Mass, Design Level Post-test



96

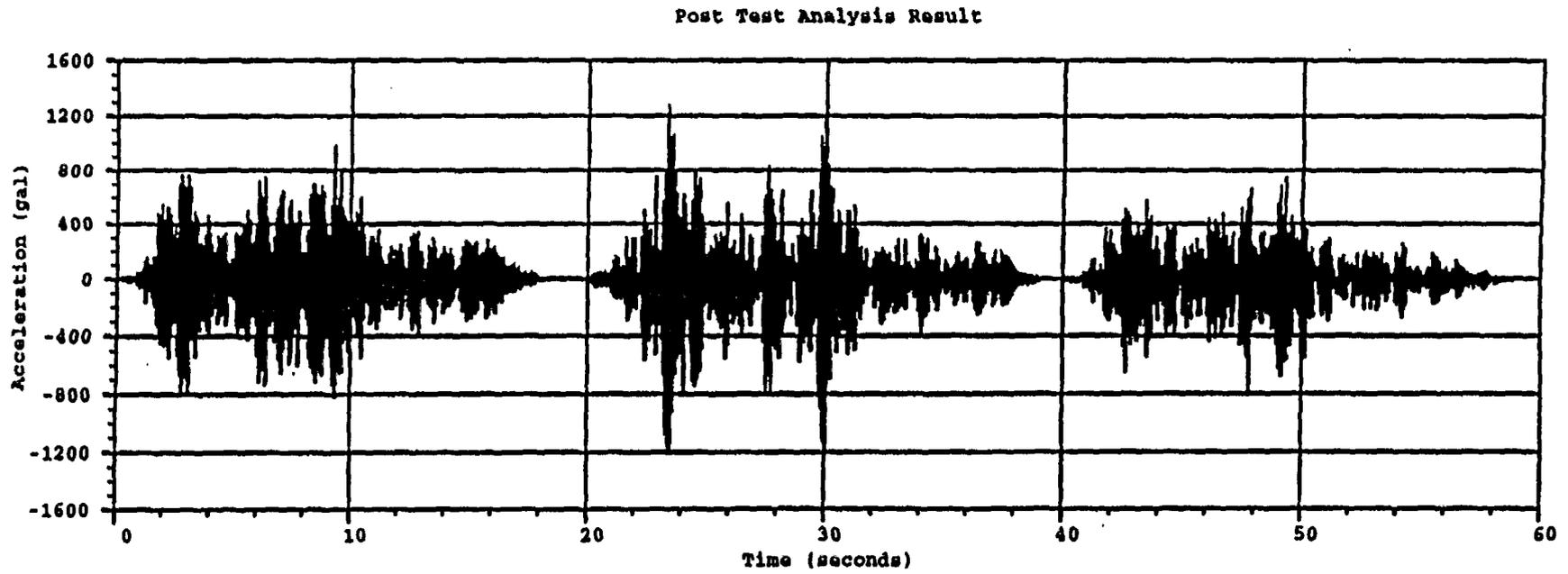
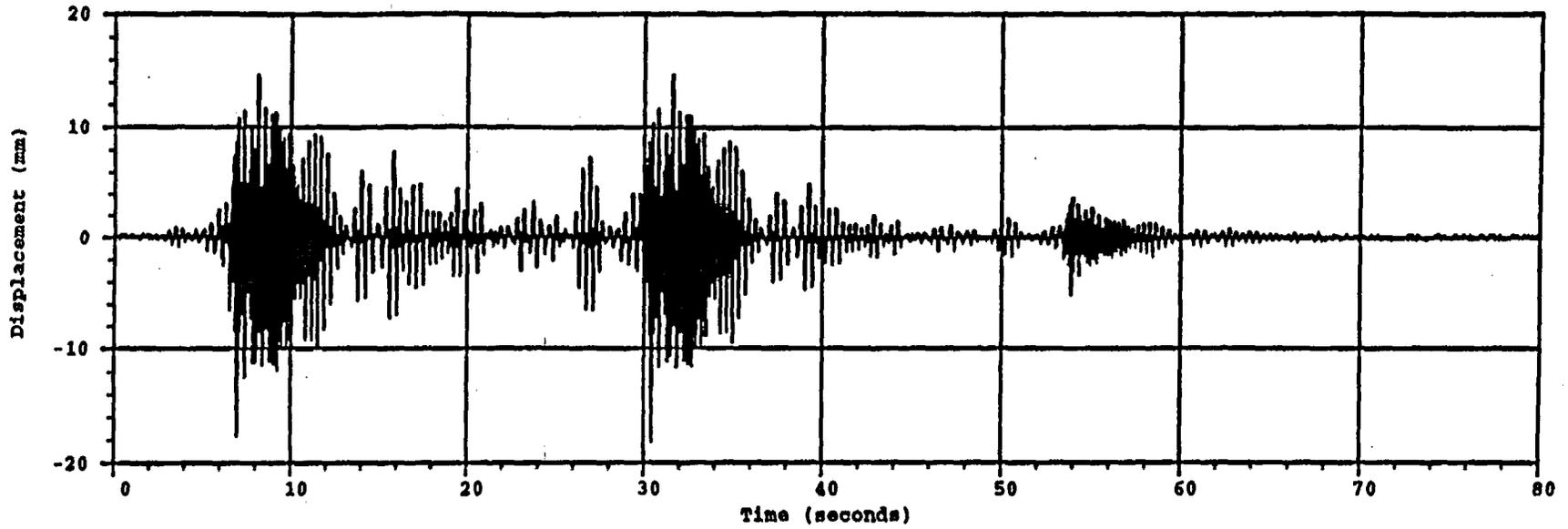


Figure 11. Horizontal Acceleration of Top Mass, Design Level Post-test

DX001 Test Result



Post Test Analysis Result

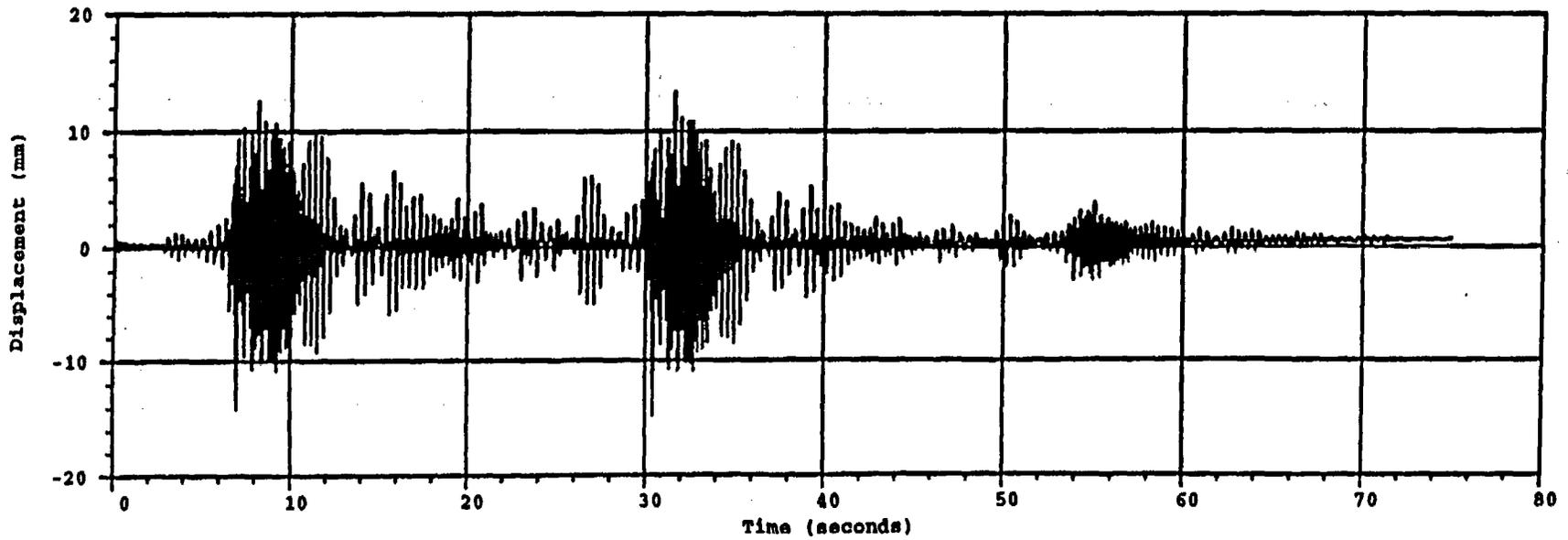
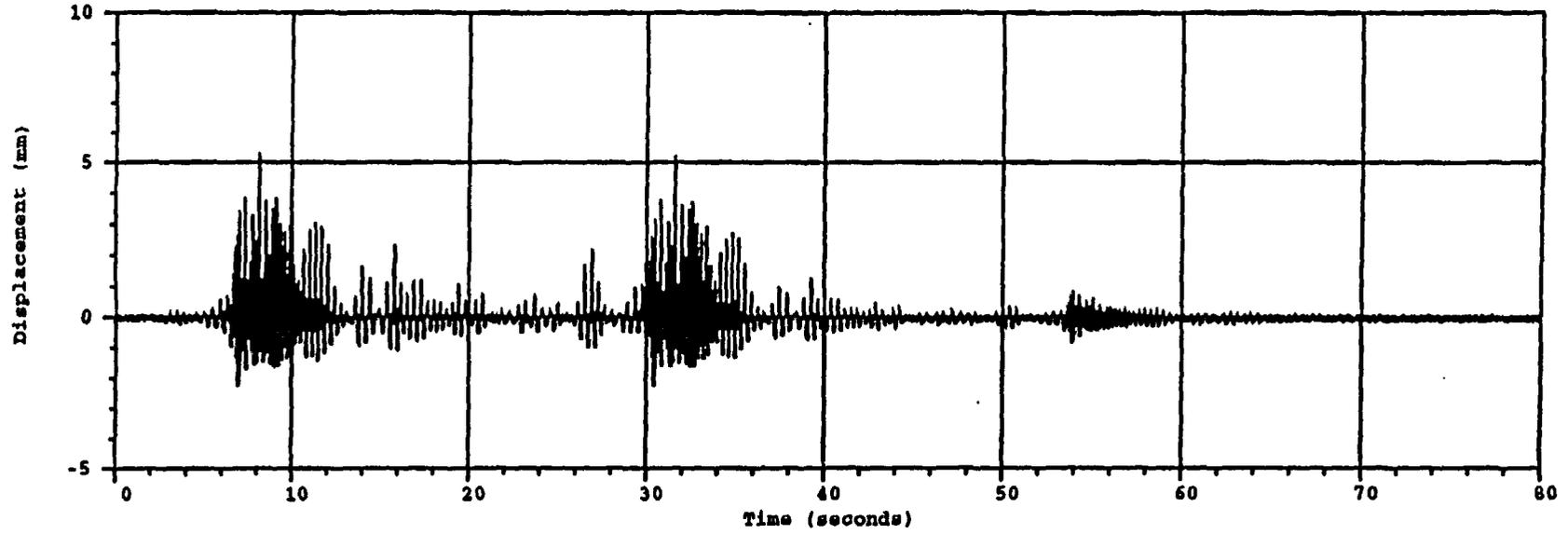


Figure 12. Relative Horizontal Displacement of Top Mass 3.3S2(H)

Test Result: Relative Vertical Displacement at 90 Deg.



Post Test Analysis Result

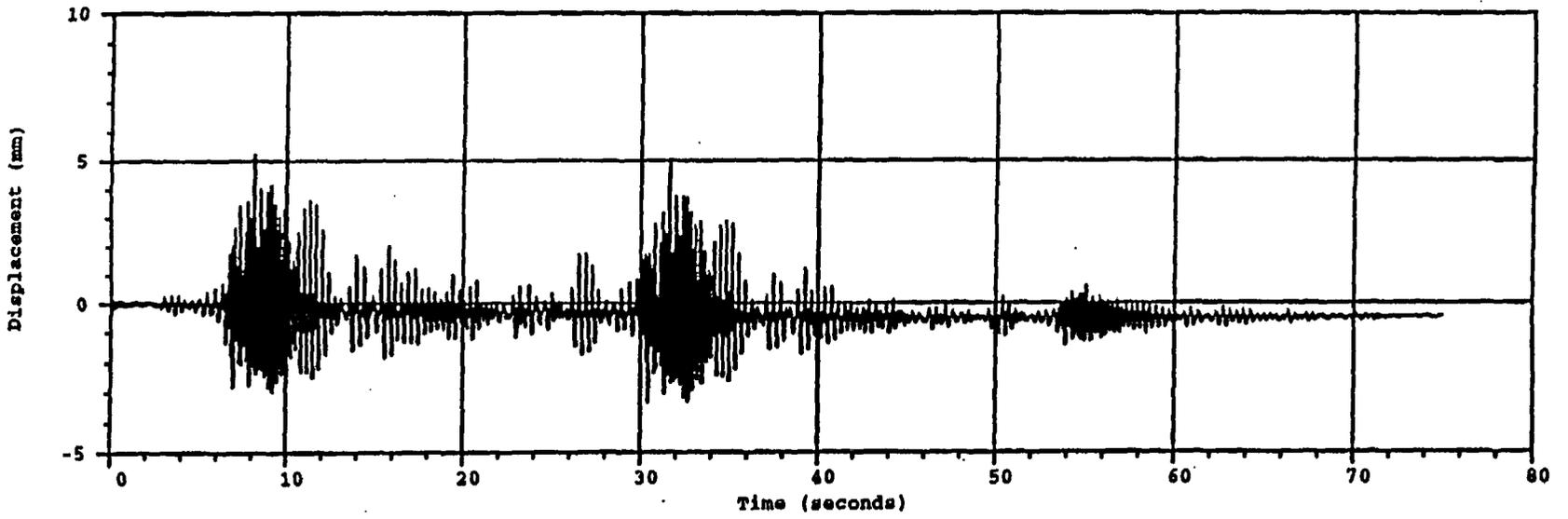
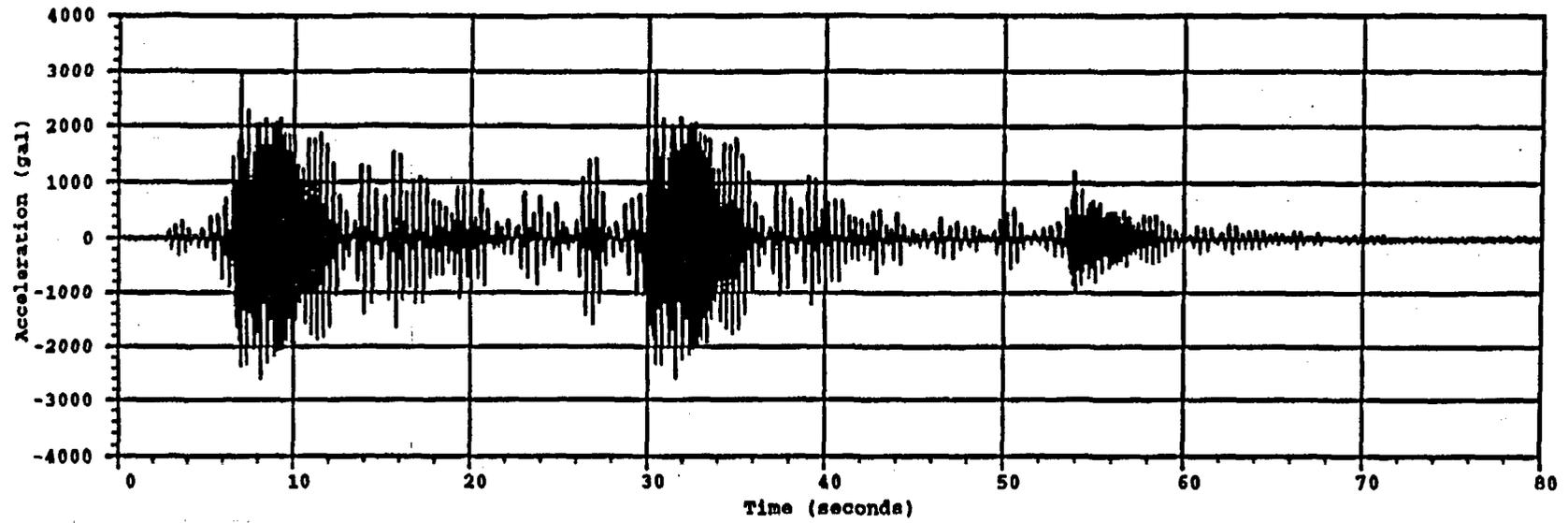


Figure 13. Relative Vertical Displacement of Top Mass at 90 Degrees 3.3S2(H)

AX001 Test Result



Post Test Analysis Result

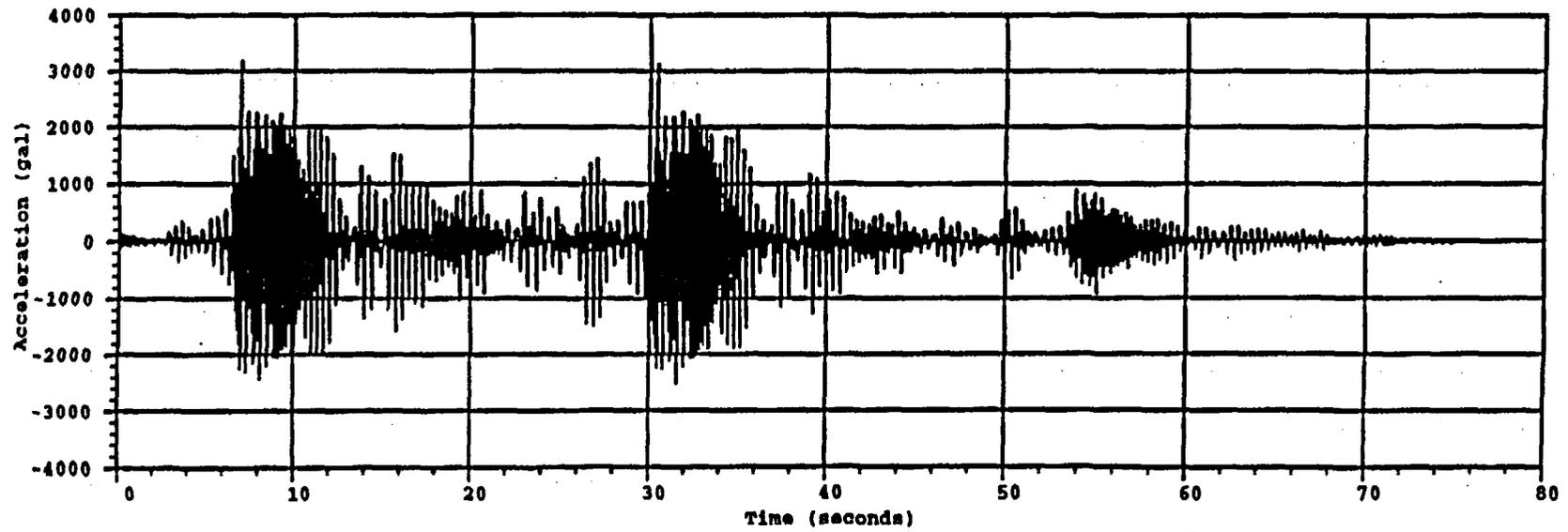
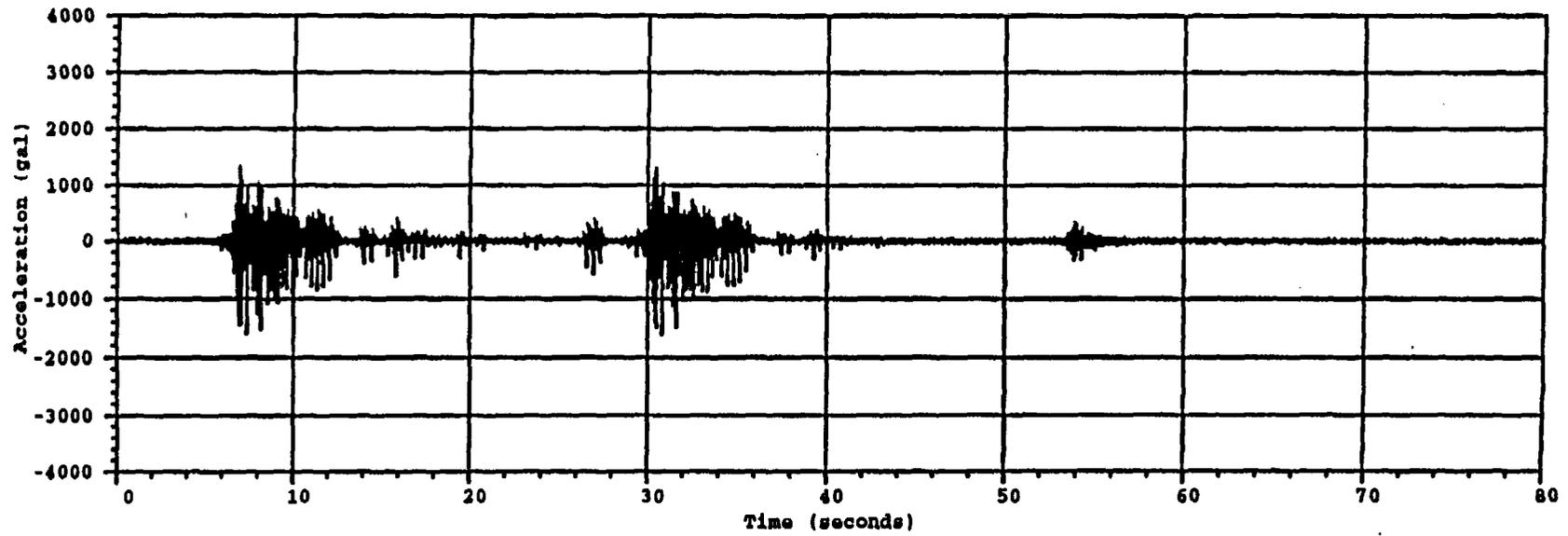


Figure 14. Horizontal Acceleration of Top Mass, 3.3S2(H)

AZ091 Test Result



Post Test Analysis Result

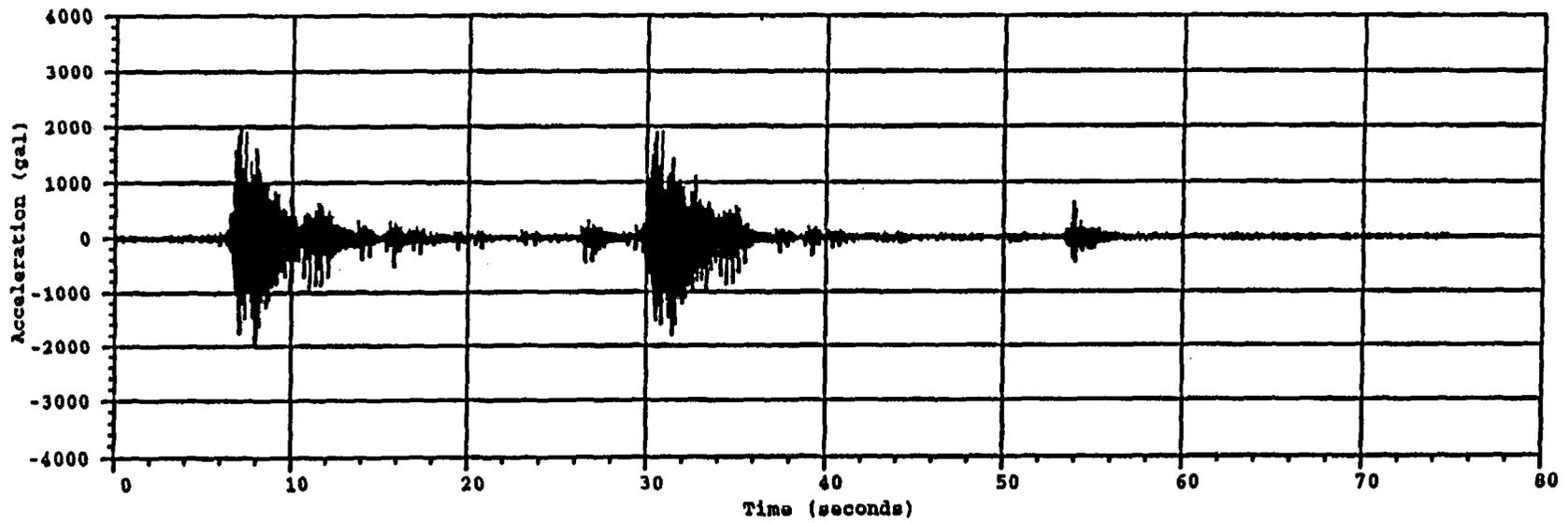
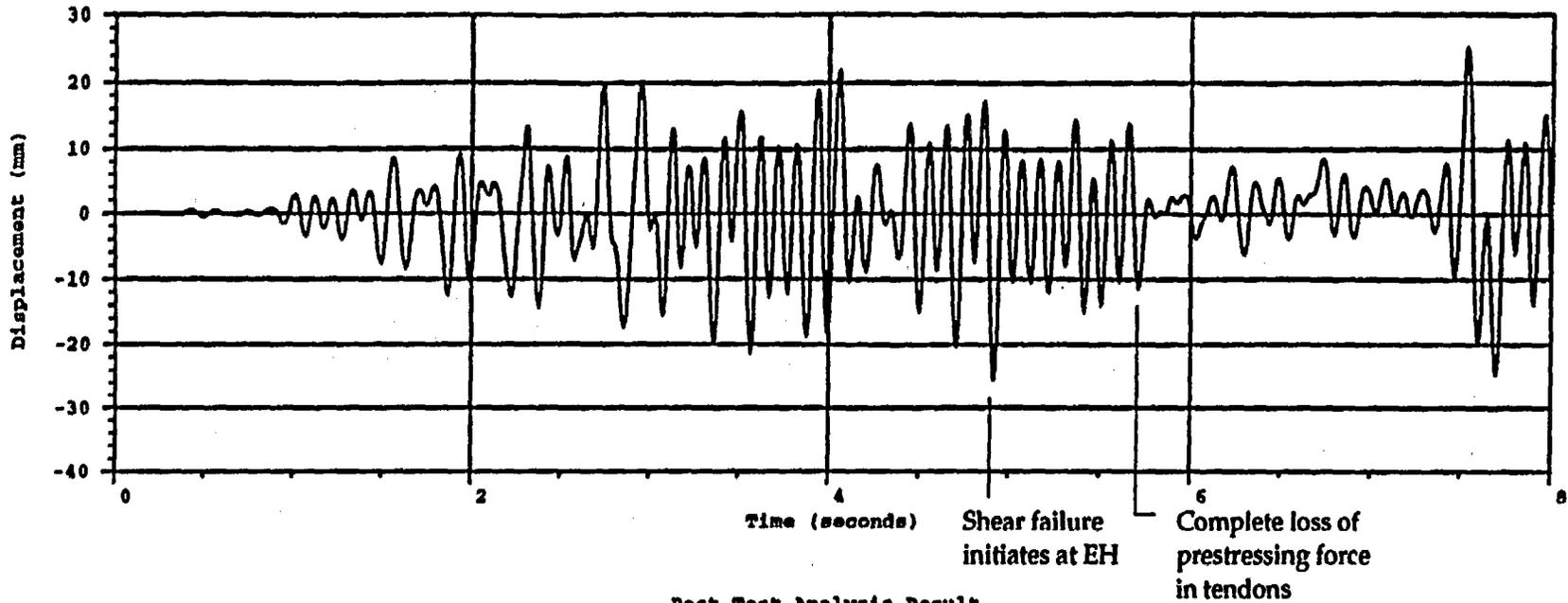


Figure 15. Vertical Acceleration of Top Mass at 90 Degrees, 3.3S2(H)

DX001 Test Result



Post Test Analysis Result

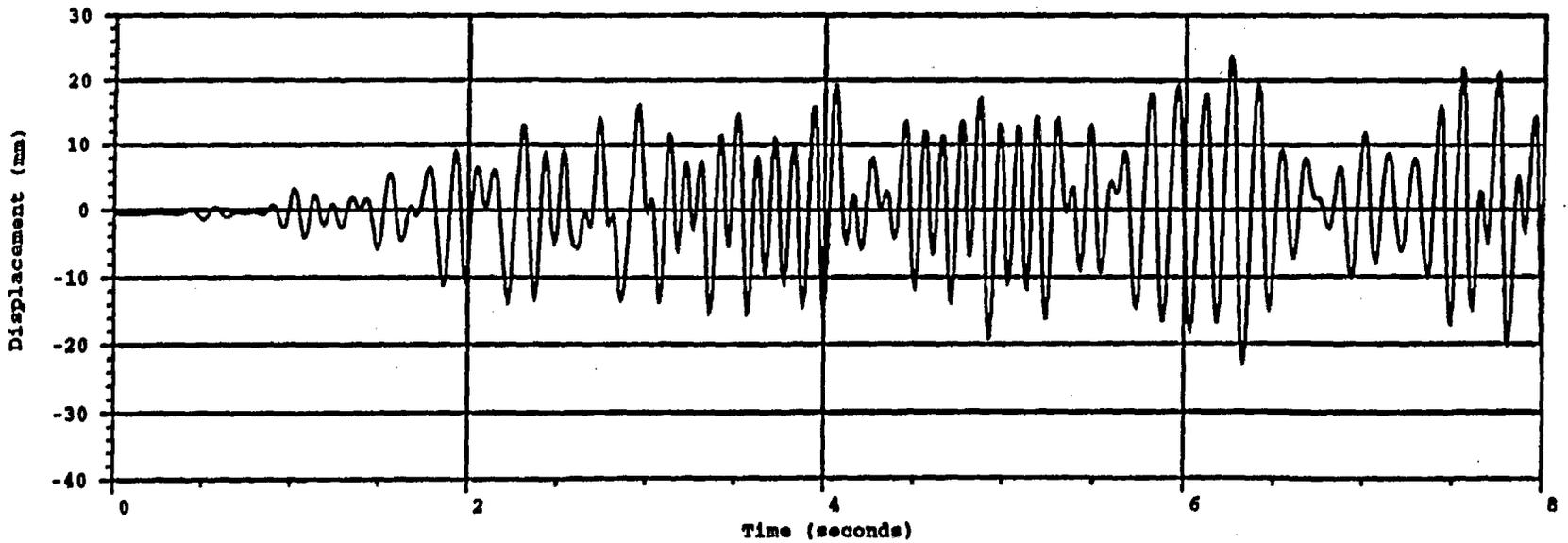


Figure 16. Relative Horizontal Displacement of Top Mass, 5S2(H)

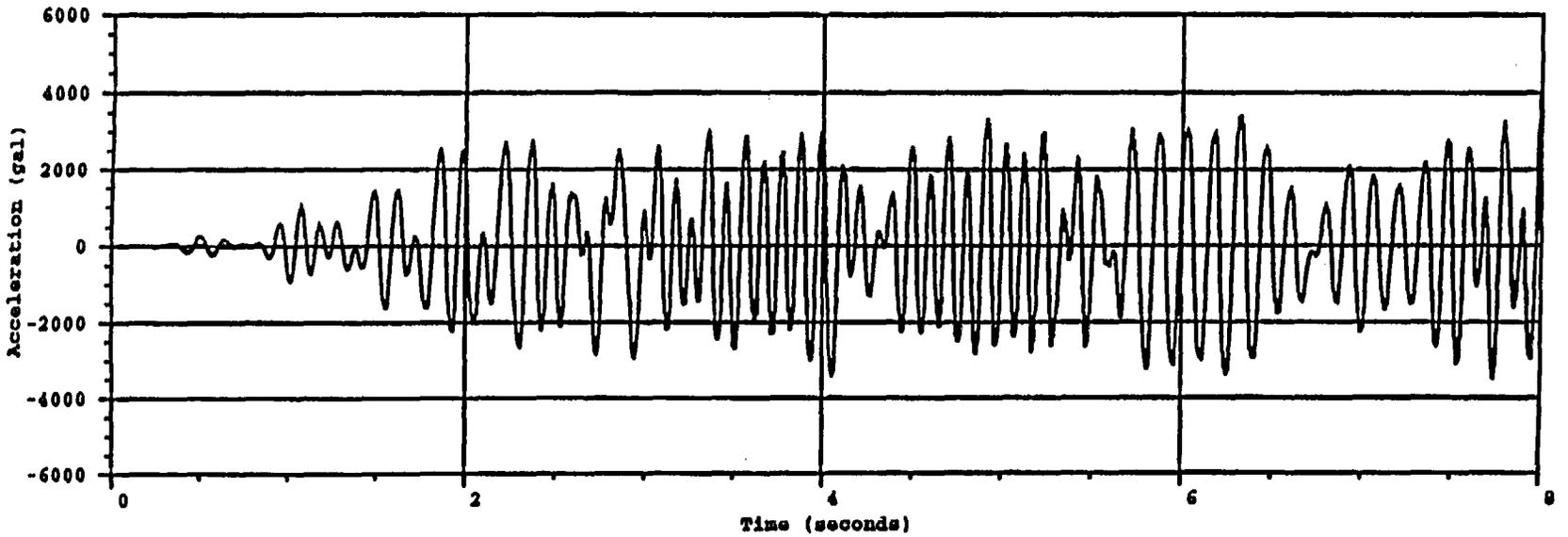
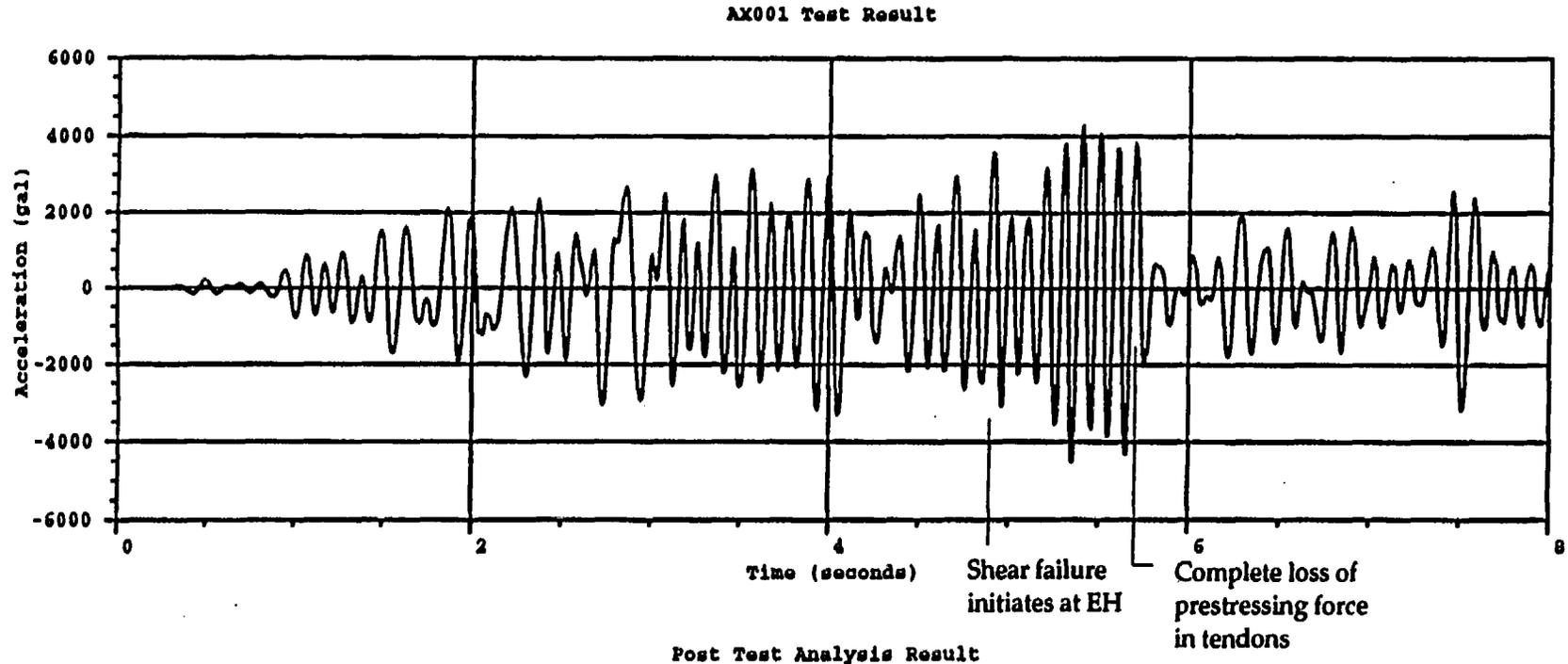
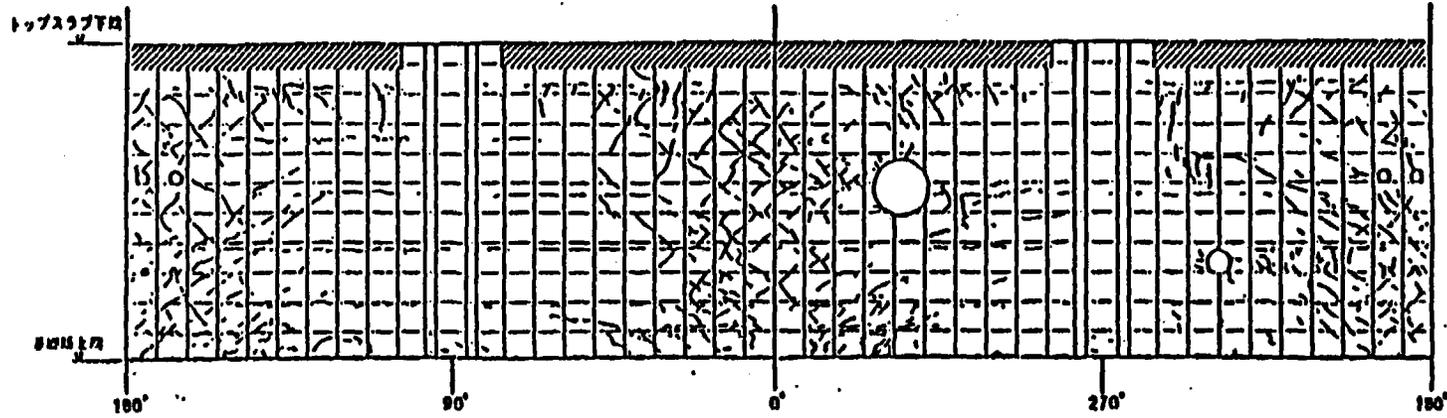


Figure 17. Horizontal Acceleration of Top Mass, 5S2(H)

Cracking Patterns After 2S2(H)



Damage After Failure in 5S2(H) Test

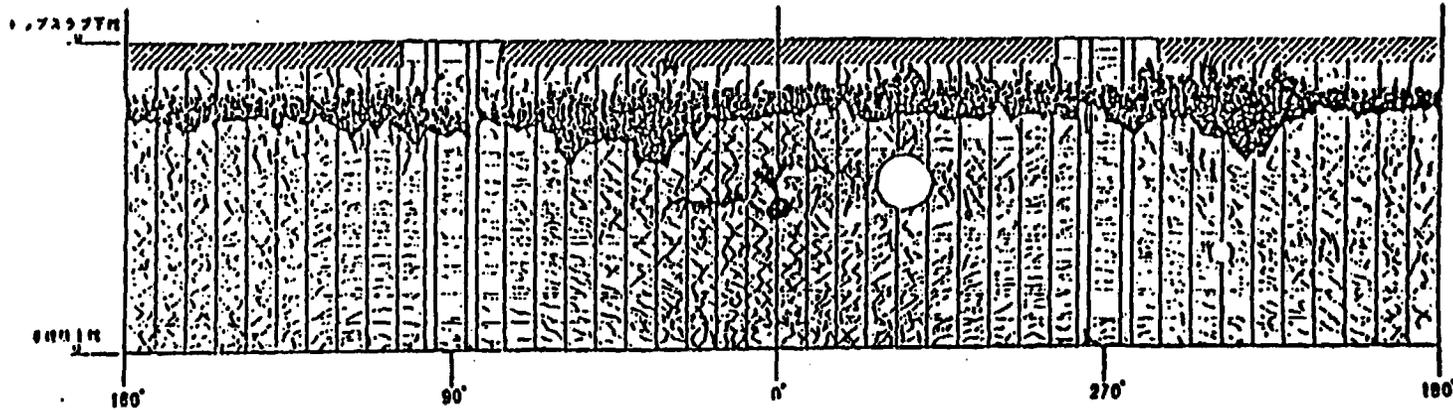


Figure 18. Cracking and Compressive Damage Exhibited in PCCV Failure Level Tests

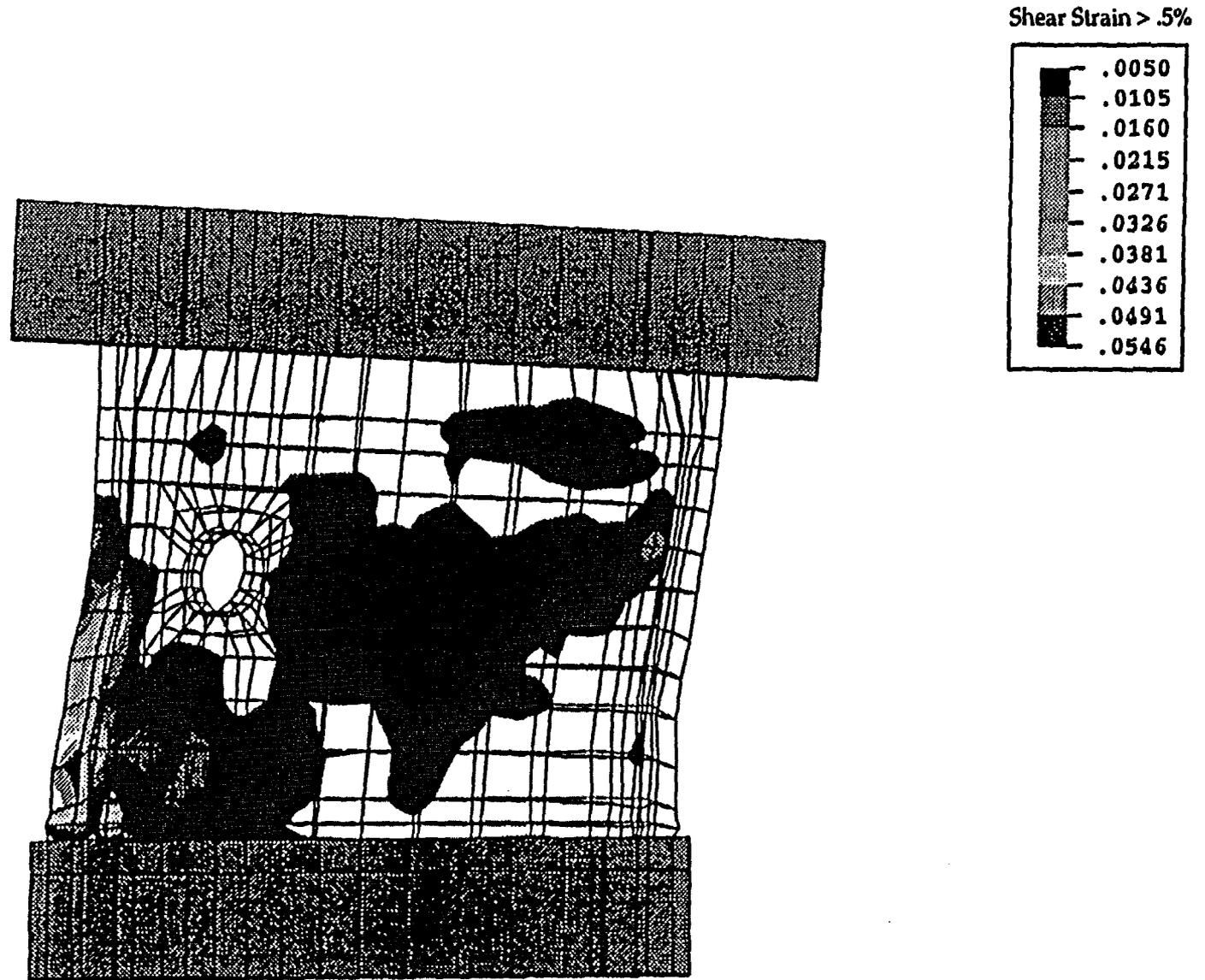


Figure 19. Shear Strains in PCCV Wall at Failure for 5.0S2(H) Test

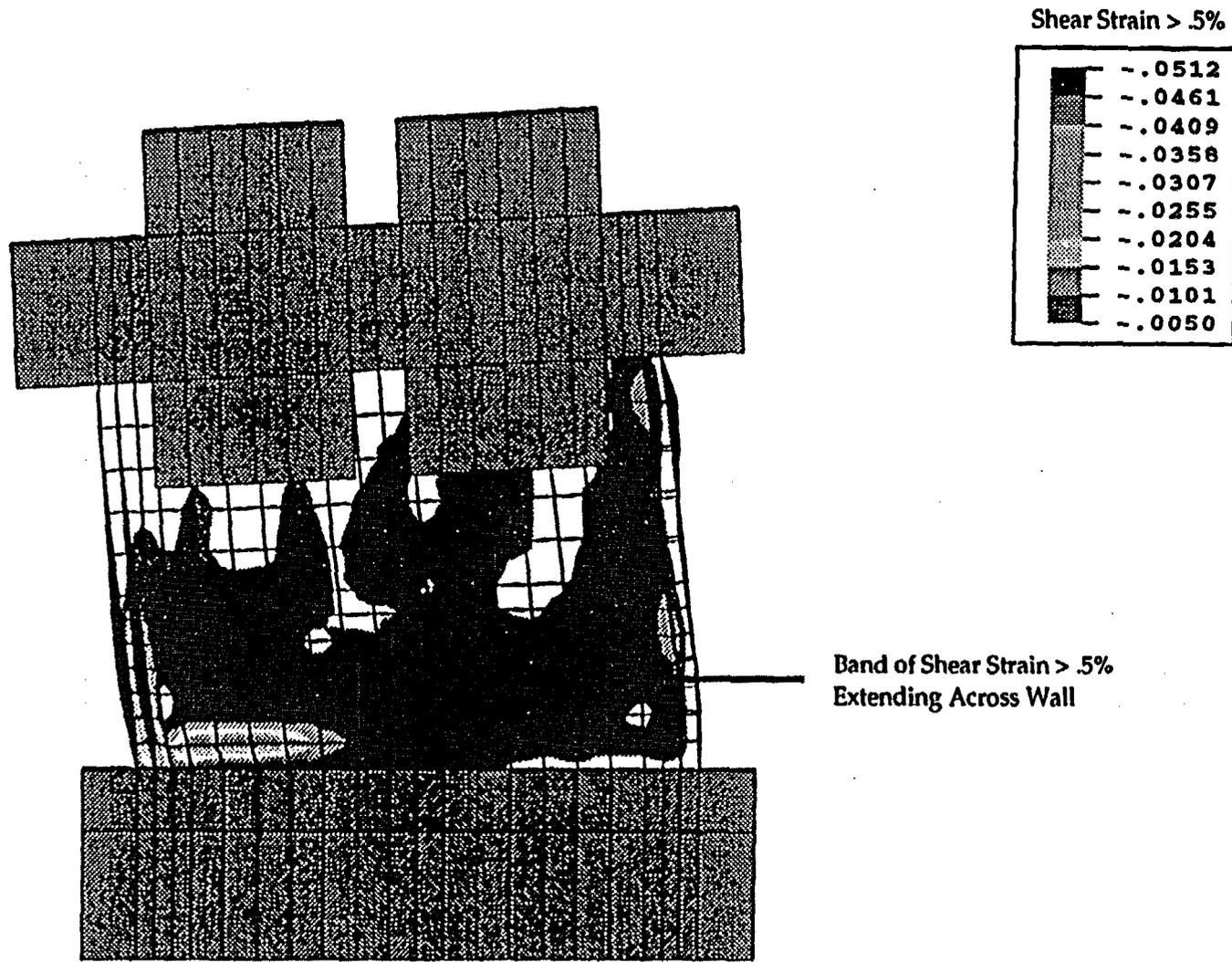


Figure 20. Shear Strains in NUPEC Shear Wall Test at Time of Failure

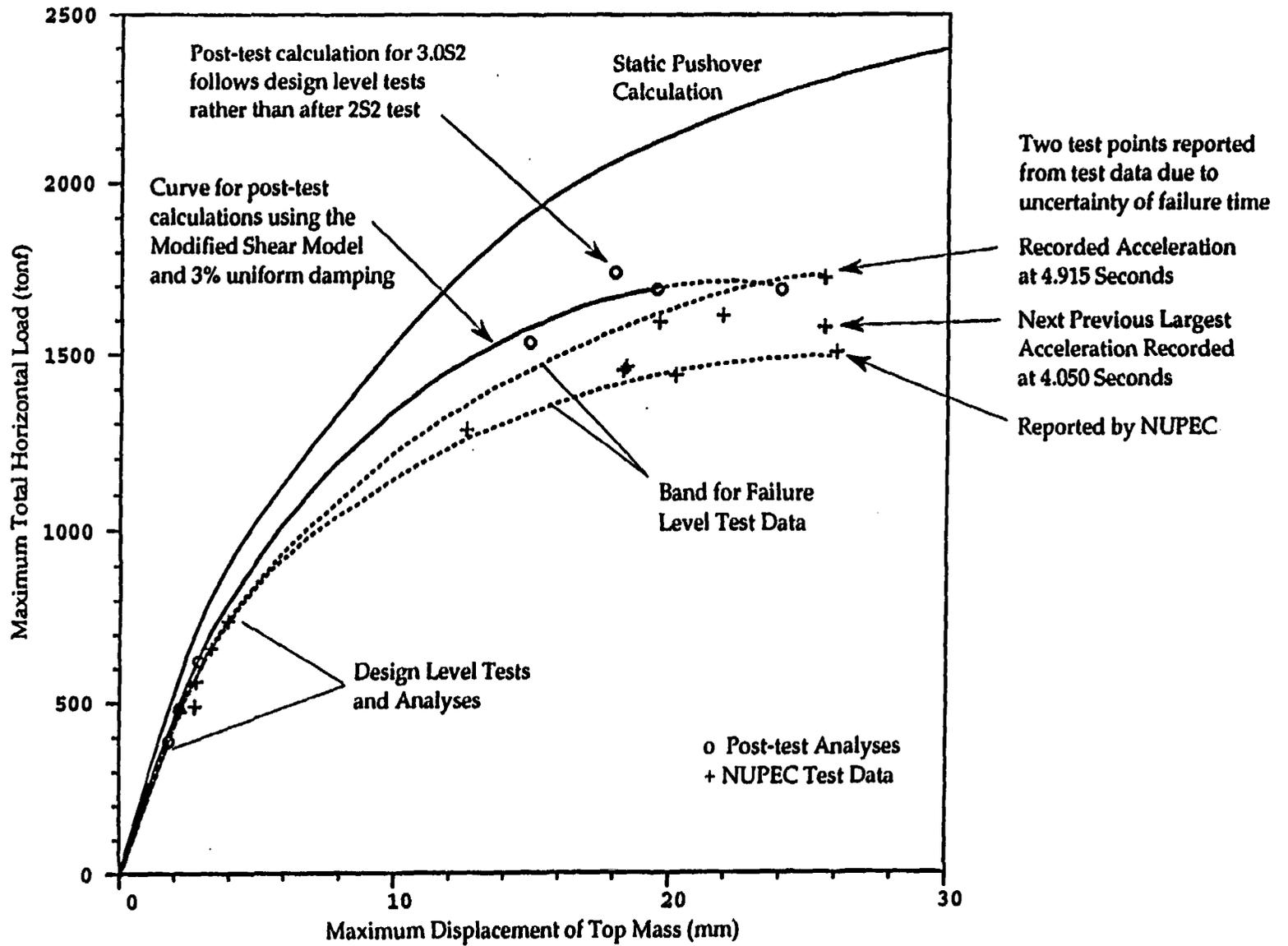


Figure 21. Summary of Peak Horizontal Response for Post-test Calculations

Steel Containment Vessel Model Test: Results and Posttest Analysis¹

Vincent K. Luk, John S. Ludwigsen, and Michael F. Hessheimer
Sandia National Laboratories, Albuquerque, NM, USA

Kuniaki Komine and Masaki Iriyama
Nuclear Power Engineering Corporation, Tokyo, Japan

Tomoyuki Matsumoto
Hitachi Ltd., Hitachi-shi, Ibaraki-ken, Japan

James F. Costello
United States Nuclear Regulatory Commission, Washington, DC, USA

ABSTRACT

Two static, pneumatic overpressurization tests of scale models of nuclear containment structures at ambient temperature are being conducted by Sandia National Laboratories for the Nuclear Power Engineering Corporation of Japan and the U. S. Nuclear Regulatory Commission. The joint research program consists of testing two models: a steel containment vessel (SCV) model and a prestressed concrete containment vessel (PCCV) model.

This paper summarizes the conduct of test of the SCV model, which is a mixed-scaled model (1:10 in geometry and 1:4 in thickness) of an Improved Boiling Water Reactor (BWR) Mark II containment, and posttest activities. A concentric steel shell, identified as the contact structure, was installed over the SCV model prior to the test to represent some of the structural characteristics of the reactor shield building in the actual plant. The SCV model and the contact structure were instrumented with strain gages and displacement transducers prior to the overpressurization test, which was conducted on December 11-12, 1996 at Sandia National Laboratories. The test was terminated when a large tear developed adjacent to the equipment hatch reinforcement plate and pressure could not be maintained in the model.

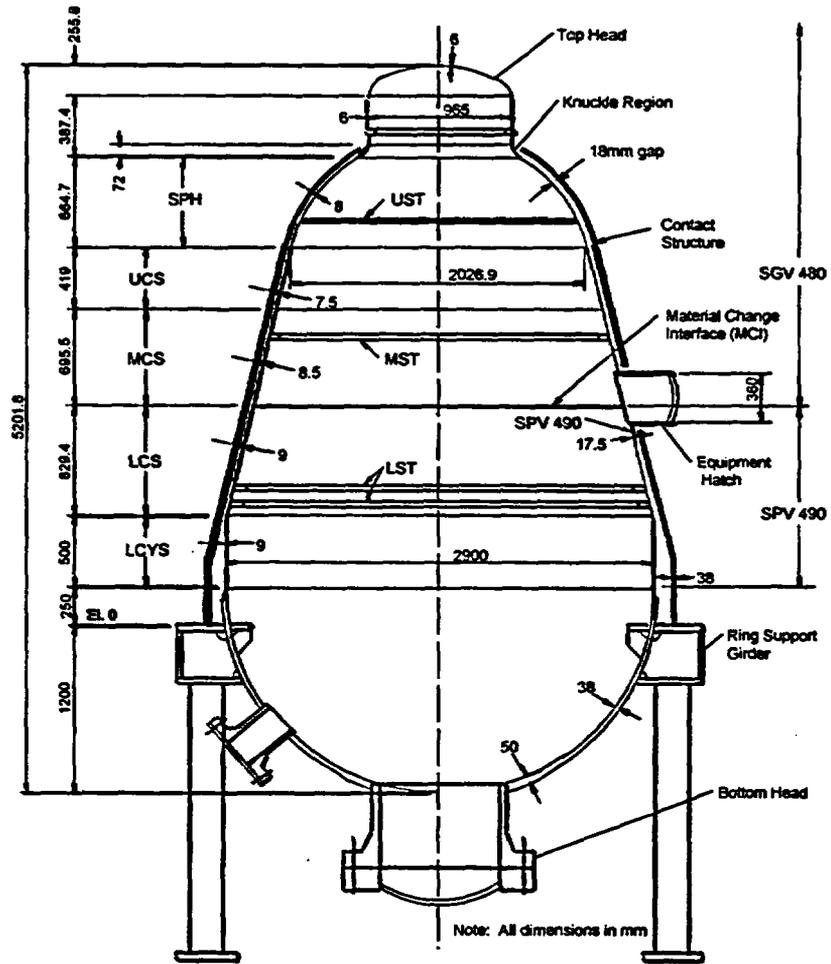
The test data are compared with the pretest analytical predictions by the sponsoring organizations and others who participated in a blind pretest prediction effort. Posttest analysis efforts focused on resolving inconsistencies between the predicted and measured free-field strains and local strain concentrations near the equipment hatch. Posttest metallurgical evaluations on specimens removed from the SCV model were also performed and the results are discussed.

INTRODUCTION

The Nuclear Power Engineering Corporation (NUPEC) of Japan and the U. S. Nuclear Regulatory Commission (NRC) are co-sponsoring a Cooperative Containment Research Program at Sandia National Laboratories. The purpose of the program is to investigate the response of representative models of

¹ This work is jointly sponsored by the Nuclear Power Engineering Corporation and the US Nuclear Regulatory Commission. The work of the Nuclear Power Engineering Corporation is performed under the auspices of the Ministry of International Trade and Industry, Japan. Sandia National Laboratories is operated for the US Department of Energy under Contract Number DE-AC04-94AL85000.

nuclear containment structures to pressure loading beyond the design basis accident and to compare analytical predictions with measured behavior. This is accomplished by conducting static, pneumatic overpressurization tests of scale models at ambient temperature. This paper describes the conduct and results of the high pressure test of the SCV model.



Nomenclature:

Location Designation

THD
 KNU
 SPH
 UST
 UCS
 MST
 MCS
 MCI
 LCS
 LST
 LCYS

Description

top head
 knuckle
 spherical shell
 upper stiffener
 upper conical shell
 middle stiffener
 middle conical shell
 material change interface
 lower conical shell
 lower stiffeners
 lower cylindrical shell

Figure 1. Elevation view of the SCV/CS assembly

MODEL DESCRIPTION

The SCV model is representative of the steel containment vessel of Improved Mark II Boiling Water Reactor plants in Japan. The geometric scale is 1:10. Since it was desired to use the same materials for the fabrication of the model as are used in the construction of the actual plants, the scale on the wall thickness was set at 1:4. The portion of the model above the material change interface which is slightly below the equipment hatch centerline (see Fig. 1) was fabricated of SGV480, a mild steel, while the lower portion of the model and the equipment hatch reinforcement plate were fabricated of high strength SPV490 steel. The equipment hatch cover and top head were non-functional in the model and were welded shut. Whereas the design pressure of the prototype containment is 0.31 MPa (45 psig), the scaled design pressure, P_{ds} , for this mixed scale model is 0.78 MPa (113 psig).

The model was fabricated at the Hitachi Works in Japan and shipped to Sandia National Laboratories in the United States for instrumentation and testing. After delivery to Sandia, a 38 mm thick steel (ASTM SA516 Grade 70) contact structure (CS) was installed over the SCV model prior to testing to represent some features of the reactor shield building in the actual plant. A nominal gap of 18 mm was maintained between the SCV model and the CS. A schematic of the SCV/CS assembly is shown in Fig. 1. Instrumentation of the model consisted of over 800 channels of data, including strain gages, displacement transducers, temperature and pressure sensors, acoustic emission device as well as video monitoring.

TEST OBJECTIVES

The objectives of the SCV model test are:

1. to provide experimental data for checking the capabilities of analytical methods to simulate the pressure response of a steel containment well into the inelastic range and after making contact with the CS,
2. to investigate the failure mode(s) of the SCV model, and
3. to provide experimental data useful for the evaluation of actual steel containment structures.

PRETEST ANALYSIS

Pretest finite element analyses were performed to predict the behavior of the model, guide the placement of instrumentation, and identify and evaluate failure modes. Details of the pretest analysis are provided in a NUREG report (Porter, et al., 1996). In addition to predicting the global response of the model, these analyses, based on the design configuration, predicted high strains in the shell surrounding the equipment hatch reinforcement plate, with the highest strains in the lower strength SGV480 shell. Generalized contact between the SCV model and the CS was predicted to occur around 4.2 MPa (600 psig). The pretest analysis predicted that the most likely failure mode was a local ductile failure at a locally thinned area (detected in a pretest inspection) in the SPV490 shell adjacent to the equipment hatch at a pressure of 4.5 MPa (650 psig). These predictions were qualified by uncertainties about the as-built configuration of the model.

In addition to the pretest analysis performed by Sandia, several organizations participated in a blind, pretest prediction exercise, euphemistically referred to as a 'Round Robin' analysis. Each participant was provided with the design and as-built information about the SCV model and the CS and was asked to

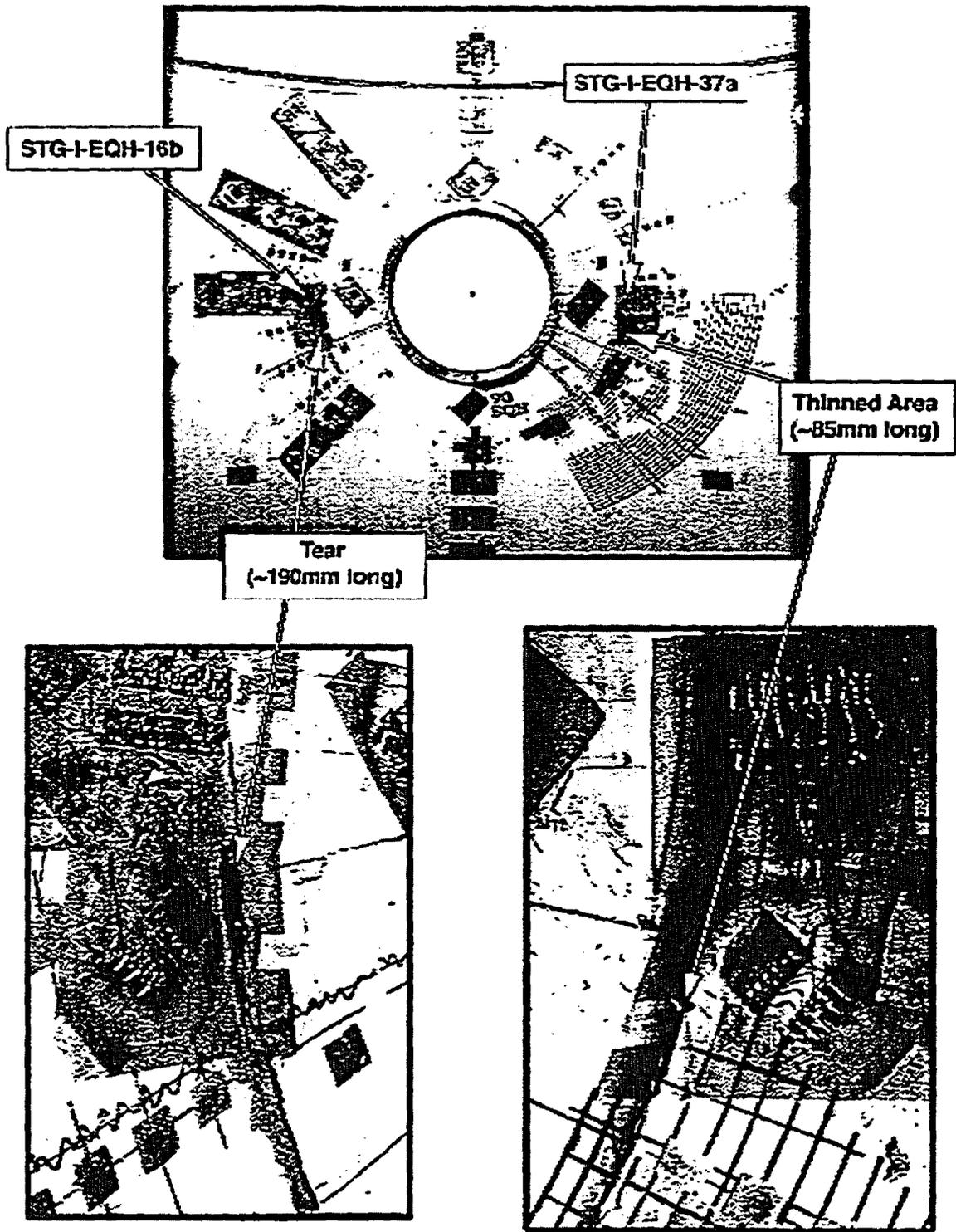


Figure 2. Interior posttest view of the equipment hatch

predict the response (strains or displacements) of the model at 43 locations as well as predicting the most likely failure mode, location and pressure. These results were compiled and published prior to the high pressure test (Luk, et al., 1996). A number of the participants met to discuss the pretest analysis results in October 1996. In general, there was a fair agreement between the independent calculations; however, there was some disagreement in the interpretation of the analysis results relative to predicting the model failure.

These pretest analyses will be discussed in more detail in conjunction with the description of the test results.

HIGH PRESSURE TEST

The high pressure test of the SCV model was conducted on December 11-12, 1996, at Sandia National Laboratories. The conduct of the test is described in a SMiRT paper and a NUREG report (Luk, et al., 1997, Luk, et al., 1998). Briefly, after approximately sixteen and a half hours of continuous, monotonic pressurization using nitrogen gas, the test was terminated when a tear developed near the equipment hatch reinforcement plate at a pressure of 4.66 MPa (676 psig) or roughly six times the design pressure. Rapid venting of the model was observed and the pressurization system, operating at capacity (1300 scfm), was unable to maintain pressure in the model.

Posttest visual inspection of the interior of the model revealed a large tear, approximately 190 mm long, adjacent to the weld seam at the edge of the equipment hatch reinforcement plate (Fig. 2). The tear appears to have initiated at a point roughly 30 mm below the material change interface (around 8 o'clock when viewed from the inside) in the higher strength SPV490 shell, and propagated in both directions along the weld seam before it stopped. Interestingly, while the right hand side of the equipment hatch (from inside view of the model) did not tear, significant necking was observed at a location symmetric with the tear (Fig. 3).

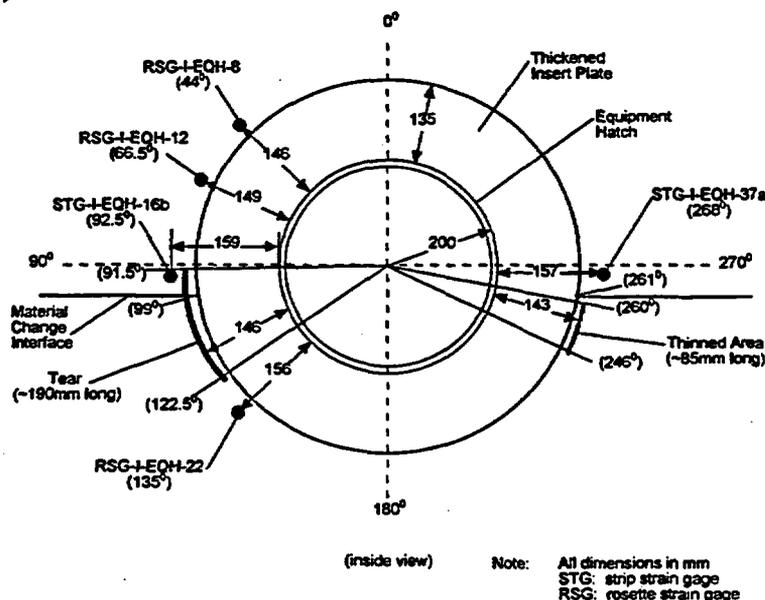
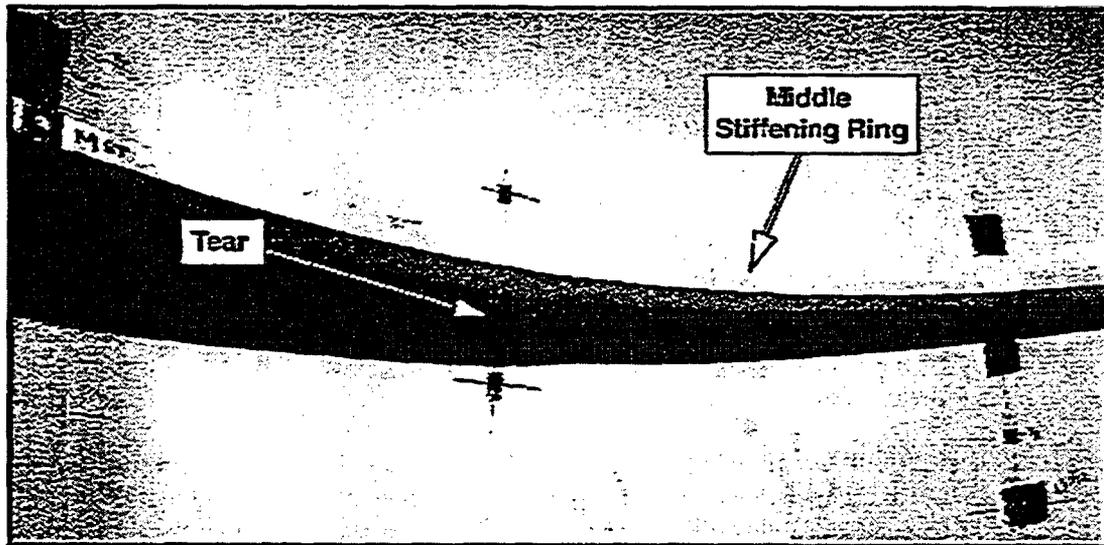
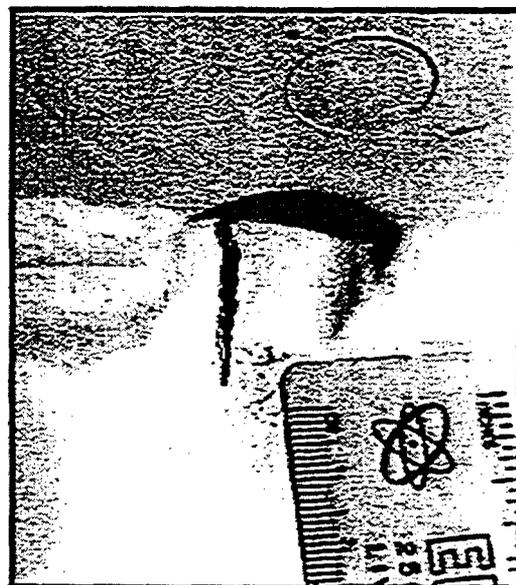


Figure 3. Interior elevation of the equipment hatch

In addition, a small meridional tear, approximately 85 mm long, was found in a vertical weld (at an azimuth angle of 201°) underneath a semi-circular weld relief opening at the middle stiffening ring (elevation of 2100 mm above the ring support girder) (Fig. 4). It appears that this small tear might have occurred first but did not grow and the pressurization system was able to compensate for any leakage through this tear. This tear also had a counterpart at a similar, diametrically opposed detail. While no tear developed at this location, necking in the weld was observed.



Above



Below

Figure 4. Posttest view of tear at middle stiffening ring

After the initial inspection of the interior of the model, the contact structure was removed to allow inspection of the exterior of the model. In addition to the observations noted above, visual inspection revealed evidence of the pattern of contact between the model and the CS in the form of crushed instrumentation lead wires and transfer of mill markings from the interior of the CS. In addition, concentrated crack patterns in the paint indicated that global strains in the higher strength SPV490 shell were concentrated at the vertical weld seams (Fig. 5).

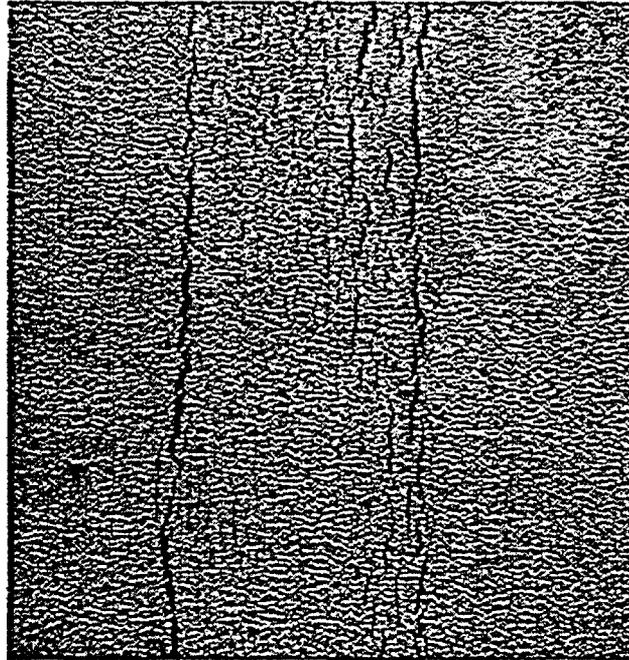


Figure 5. Exterior posttest view of vertical weld seam in lower conical shell section (SPV490)

TEST RESULTS AND COMPARISON WITH PRETEST PREDICTIONS

More than 97% of the instruments survived the high pressure test. The failed gages, which consisted primarily of those on the exterior of the model, were damaged when the model made contact with the CS. The raw strain data was corrected to compensate for temperature variations and cross-axis strains and the displacement data was corrected to account for any movement of the center support column to which the displacement transducers were anchored. The complete data record is included in the SCV Test Report (Luk, et al., 1998). A brief summary of the test data follows.

Local Response Adjacent to the Equipment Hatch

An extensive array of single element, strip and rosette strain gages was installed around the equipment hatch to capture the local strain distribution. Figure 3 shows the locations of a few selected strain gages around the equipment hatch viewed from the inside of the model. A strip gage (STG-I-EQH-16b) installed adjacent to the upper end of the large tear registered a maximum hoop strain of 4.2 % and the two rosette gages (RSG-I-EQH-12a and -8a) above it had recorded maximum hoop strains of 3.7 % and 2.8 %, respectively. The rosette gage (RSG-I-EQH-22a) slightly below the lower end of the tear recorded a maximum hoop strain of 1.3 %. However, the highest hoop strain reading of 8.7 % was recorded by a

strip gage (STG-I-EQH-37a) at 3 o'clock, just above the material change interface. Figure 6 shows the strain data recorded by these gages around the equipment hatch.

While the pretest calculations predicted failure in the vicinity of the equipment hatch at pressure levels very close to the actual failure pressure, a detailed comparison of the calculated and measured strains highlights some areas of discrepancy. First, the highest measured strains occurred in the higher strength SPV490 shell, below the material change interface, rather than in the weaker SGV480 shell as predicted by the analyses on the design configuration of the model. Second, the near-field measured strains around the equipment hatch were almost double those predicted by the analysis. Finally, the locally thinned area, which was the predicted failure location in the pretest analysis, appeared to have little effect on the model response in the vicinity of the equipment hatch.

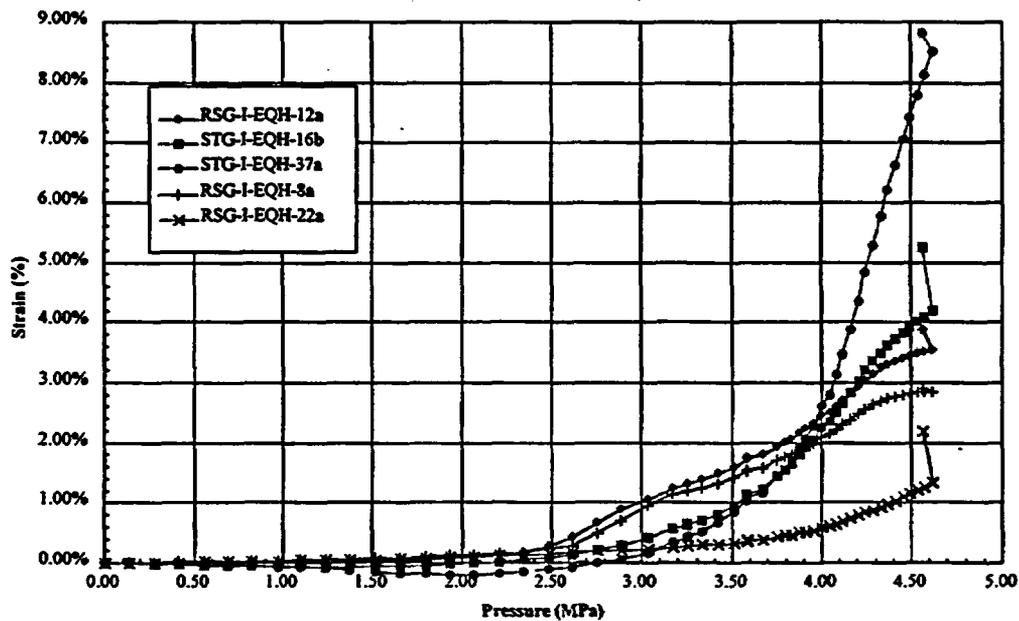


Figure 6. Hoop strains around equipment hatch

Global Response

The global response of the SCV model was monitored using free-field strain gages and an array of internal displacement transducers that measured the strains and displacements at several elevations along 4 cardinal azimuths (0°, 90°, 180°, and 270°).

Maximum free-field hoop strains ranging from 1.7 to 2.0 % were measured at 4.5 MPa (560 psig) at the upper conical shell section (Fig. 7). The hoop strains calculated from the displacement measurements ($\Delta r/r$) were consistent with the strain gage measurements at these locations.

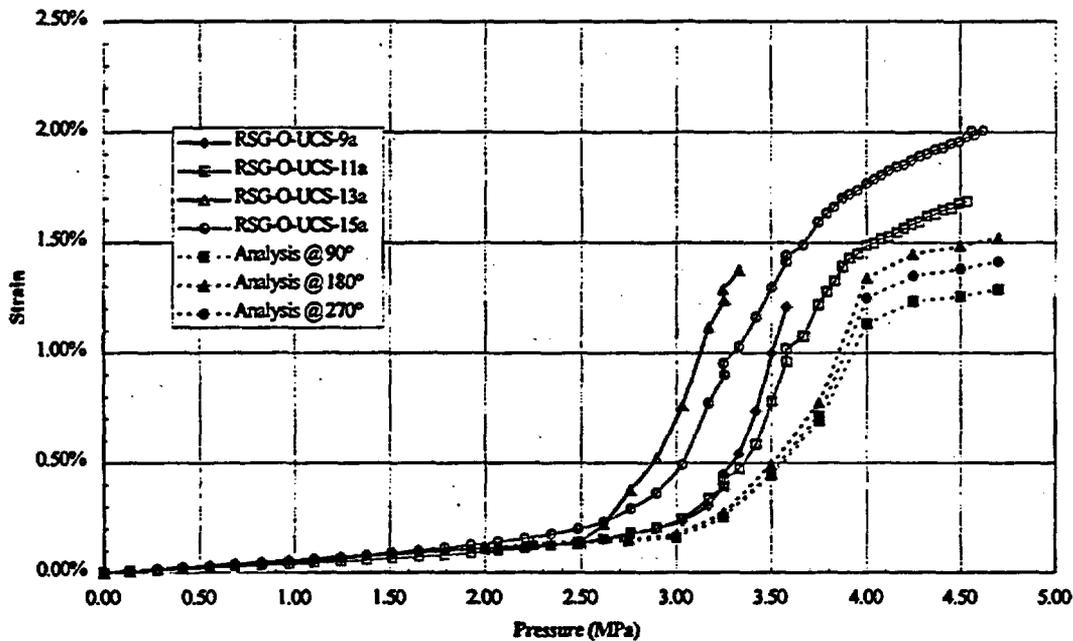


Figure 7. Hoop strains @ upper conical shell section (UCS), EL. 2536 mm

Figure 8 shows the spatial variation of displacements at the cardinal azimuths at 4.5 MPa. The displacement profiles were completed by interpolating the data recorded by the transducers at various elevations. It should be noted that the displacement pattern is fairly axisymmetric with the exception of 90°, the azimuth where the equipment hatch is located. The displacements at this azimuth in the lower conical shell section, below the material change interface, were much larger than those at the free-field azimuths (0°, 180°, and 270°). This is of particular interest in light of the fact that this area was actually displaced inward during fabrication of the model and this was also the area where the large tear occurred.

Figure 9 shows the spatial variation of displacements as a function of pressure at a representative free-field azimuth (270°). This figure indicates a disproportional increase in radial displacement of the model between 3 and 4 MPa, suggesting that the global yielding of the model might occur somewhere in this pressure interval. Observable slow-down in radial growth of the model occurred beyond 4 MPa when the model made local contact with the CS.

Additional displacement plots at the middle and upper conical shell sections as a function of pressure are shown in Figs. 10 and 11, respectively. The manner in which the plots in these two figures and Fig. 7 started to curve upward at about 2.5 MPa suggests that the onset of yielding of the model might have occurred as early as 2.5 MPa. Additionally, it can be inferred from these figures that generalized contact between the model and the CS began at pressures between 4.0 to 4.5 MPa.

Figures 7, 10, and 11 also compare the pretest analysis predictions for global strains and displacements with the test results. The most significant observation from this comparison is that the pretest

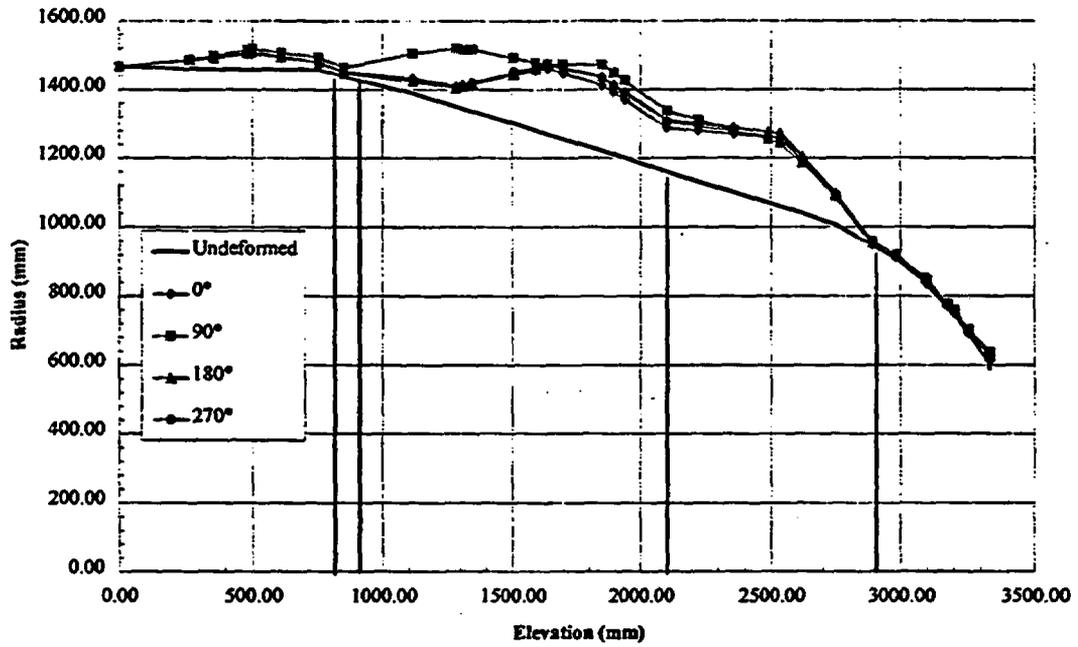


Figure 8. Displacement contours (x10) @ 4.5 MPa

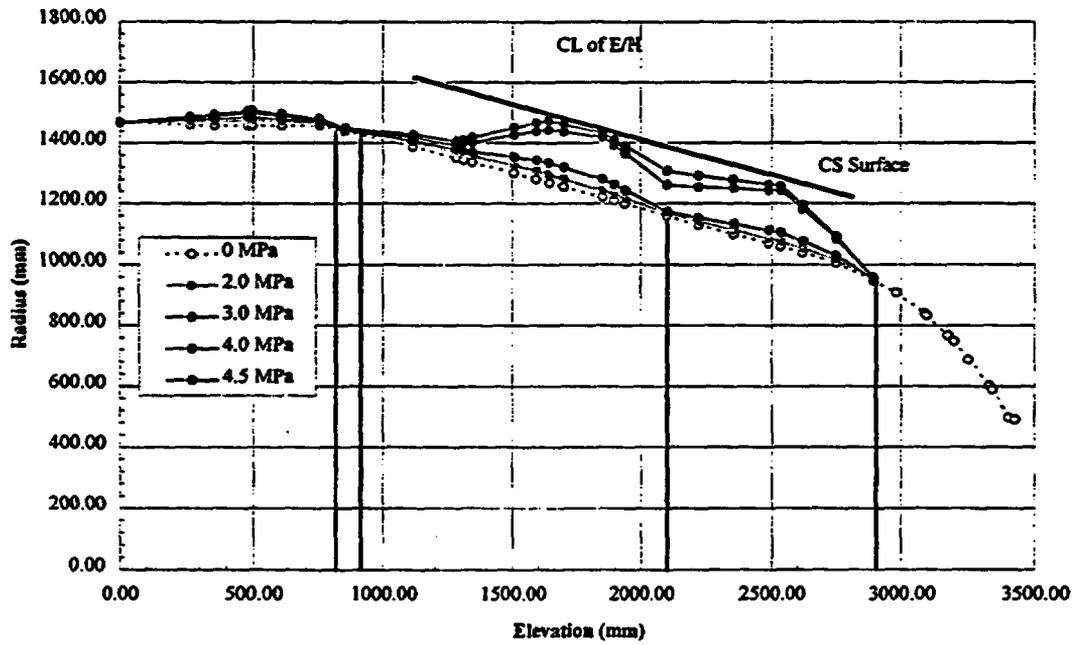


Figure 9. Displacement contours (x10) @ 270° azimuth

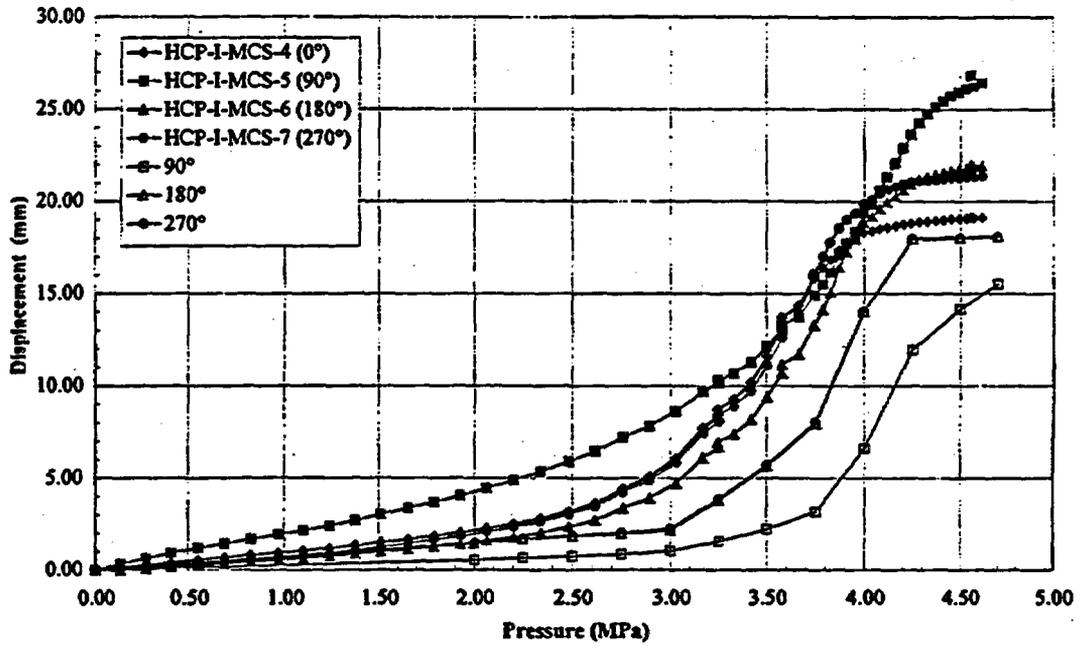


Figure 10. Radial displacements @ middle conical shell section (MCS), El. 1850 mm

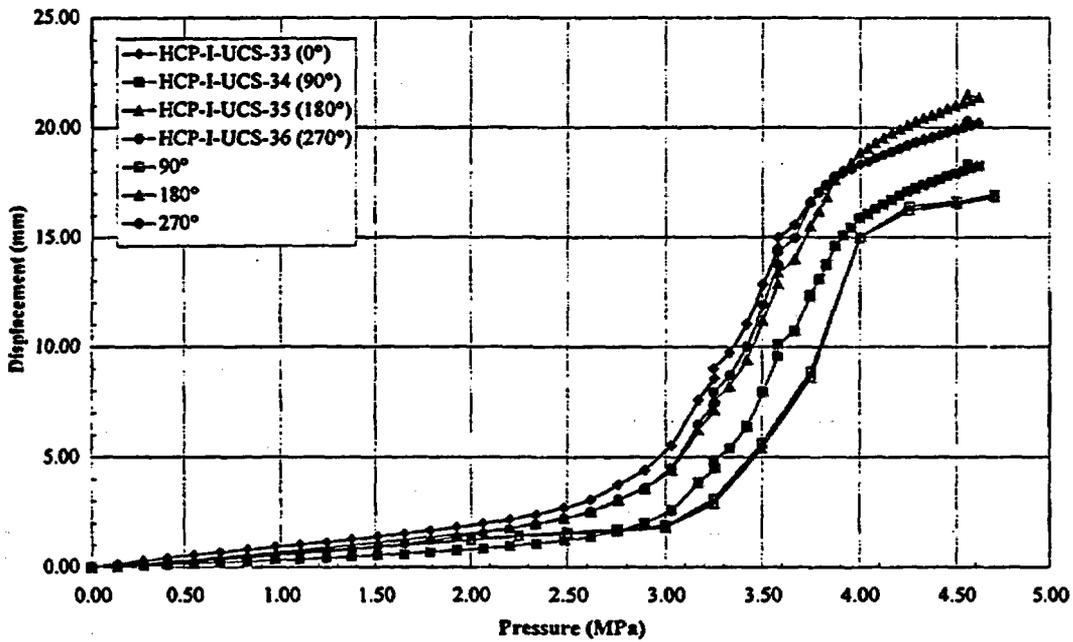


Figure 11. Radial displacements @ upper conical shell section (UCS), El. 2536 mm

calculations significantly overpredicted the pressure at which the global yielding occurred and continued to underpredict deformations and strains after yielding up to model failure. This comparison result is troubling and unexpected. Good agreement between the global response calculations and the test results was expected, based on past experiences, when uniaxial tensile test data for the actual materials used in the fabrication of the model were used to define the material properties for analytical model.

In attempting to understand the source of this disagreement, a comparison between the analytical results of several Round Robin participants and the test data was made (Fig. 12). This figure illustrates the effect of using the lower bound, average and actual results of uniaxial tensile test data to define the material properties for the model. From this comparison, it appears that the use of the lower bound material data gave the best agreement with the test results. There may be a variety of contributing factors to the discrepancy between the analytical and test results; however, this comparison highlights the sensitivity of the analytical results to relatively small variations in the material models.

One other observation from this comparison is that it appears that the effective gap was larger than the nominal gap of 18 mm used in the pretest analysis. No attempt was made to characterize the as-built gap in the pretest analysis, even though this dimension varied from 13 to 24 mm after the installation of the contact structure was complete.

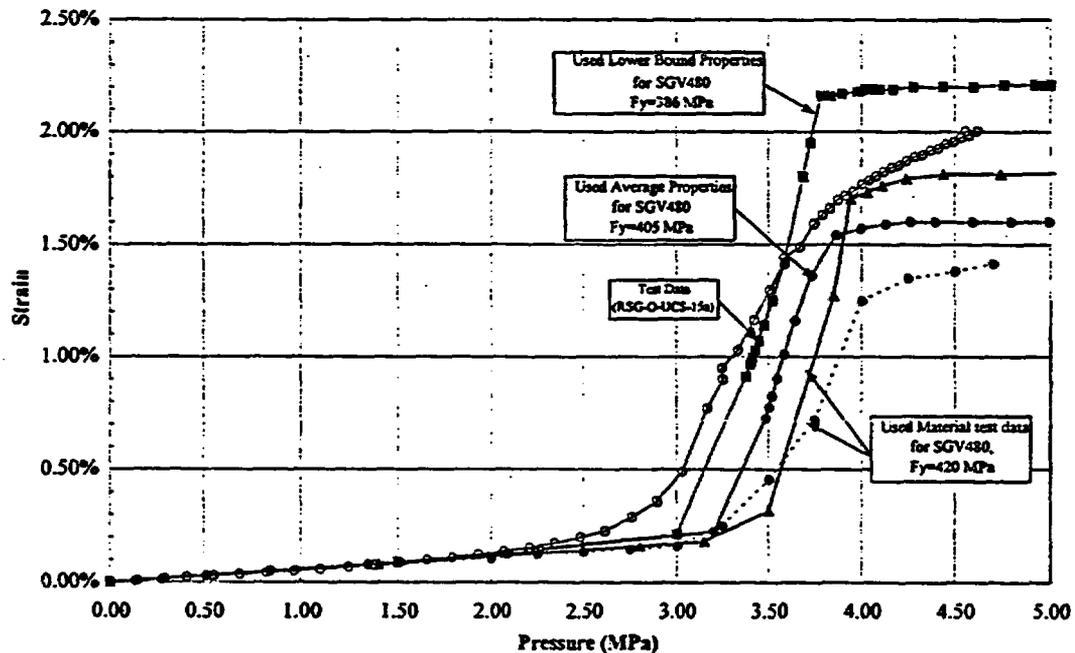


Figure 12. Hoop strains @ upper conical shell section (UCS), El. 2536 mm

Acoustic Emission Data

In addition to the strain gages and displacement transducers, twenty-four acoustic emission sensors (eighteen interior and six exterior) were installed on the model. Posttest analysis of the data collected by these sensors indicated two regions with high acoustic emissions during the test. One region located just below the equipment hatch began generating significant acoustic activity at approximately 4.25 MPa. The close proximity of this region to the equipment hatch suggests that significant material distress, leading to the large tear might have begun at this pressure. Another region had a significant increase in acoustic emissions beginning at 3.75 MPa, however, this region is not very close to the small tear, and therefore it is not clear whether the initiation of the small tear is related to this pressure.

POSTTEST INSPECTION AND EVALUATION

In addition to the posttest visual inspection described above, a detailed metallographic evaluation of the SCV model was conducted to characterize the local failure mechanisms and provide some insight into both the global and local responses of the model. The detailed evaluation and analysis are described in a SAND report (Van Den Avyle, et al., 1998).

Briefly, sections were removed from the model surrounding the tears and areas of necking or other obvious structural distress. Fractographic inspection of the failure surfaces indicated that the tearing mechanism was ductile and did not display any evidence of flaws or other defects that might have acted to initiate failure. It was therefore concluded that the model failure occurred as a result of strains exceeding the limits of the material and it should be possible to characterize failure based on the material properties of the steels.

Smaller samples were machined from the sections removed from the model and the polished cross-sections normal to the model surface were examined using a scanning electron microscope to characterize the grain structure. Hardness tests were also performed on these polished samples to look for variations in material properties. A cross-sectional view through the major tear at the equipment hatch is shown in Fig. 13.

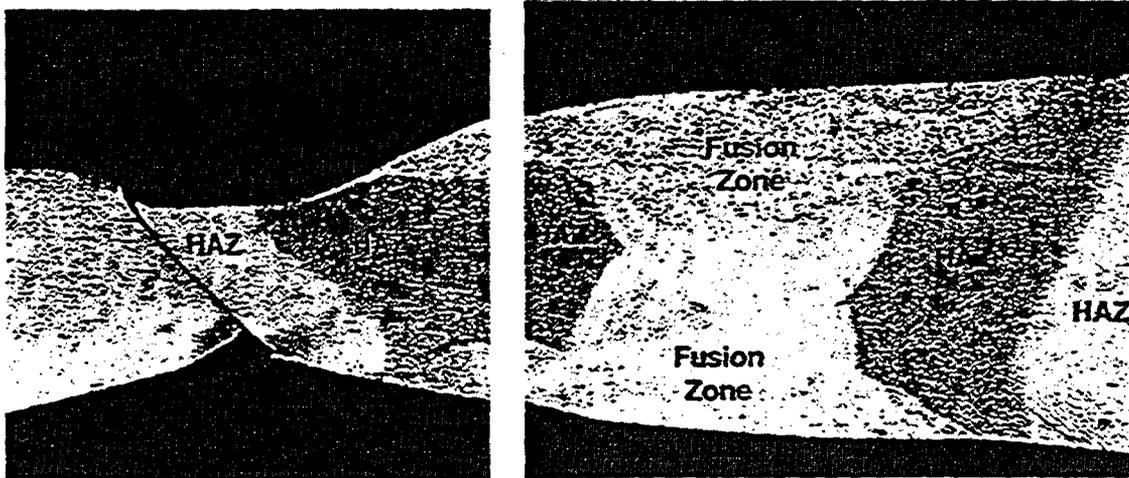


Figure 13. Cross-section through large tear @ equipment hatch

The results of these inspections revealed changes in the grain structure of the SPV490 steel in the heat affected zone (HAZ) surrounding the reinforcement plate weld and a significant reduction in the hardness of the HAZ and adjacent parent material. Based on the well-established relationships between hardness and tensile strength, these results indicate a significant reduction in tensile strength along with a corresponding, though less well defined, reduction in the yield strength of the material. These results suggest that this localized microstructural alteration and reduced hardness and strength in the HAZ of the SPV490 alloy plate may be one of the possible causes for the observed strain patterns around the equipment hatch and in the weld seams of the SPV490 shell.

POSTTEST ANALYSIS

Considering the SCV model test data, the posttest visual and metallographic evaluations of the SCV model, and the pretest analysis results, the posttest analysis effort was focused to address the observed behavior of the model and the inconsistencies between the pretest analysis results and the test data. The results of the posttest analysis are summarized in the following subsections.

Material Modeling

In an effort to address the discrepancy between the pretest analysis results and the test data of the onset of yielding in the free-field of the model, the pretest material models for the two steels were critically evaluated. As can be seen in Fig. 14, the pretest material model for SGV480 steel, based on a hardening curve-fitting scheme with an inverse hyperbolic sine curve, provided a very good representation of the measured tensile coupon test data in the high strain regions (over 20 %), but did not closely simulate the material behavior at strain levels below 5 %. For the posttest analysis, a much simpler approach was used to model the material behavior of these steels. The lower envelope of the plots of true stress versus true strain from the tensile coupon tests was used to model the plastic behavior of the materials, and the elastic portion of the stress-strain curve assumed a handbook value for the Young's modulus.

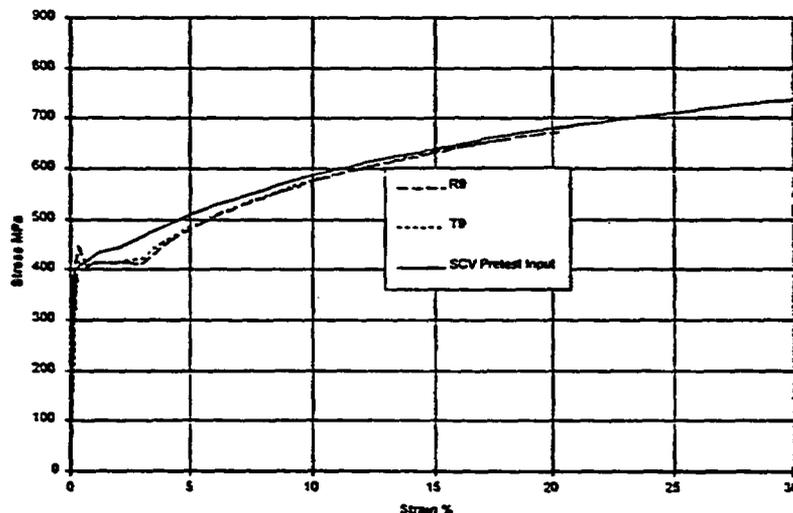


Figure 14. The pretest material model and the tensile coupon test data for 8.5 mm SGV480 steel

The posttest metallurgical evaluation results indicate that the large tear occurred inside the HAZ of SPV490 steel shell whose material strength was significantly reduced. A material model with reduced strength for SPV490 HAZ is thus needed to provide a better simulation of the strain distribution around the large tear. An approximate hardness number for the pretest SPV490 HAZ was obtained from the available hardness measurements on the posttest HAZ and base metal, and on the virgin plate material. The tensile strength of SPV490 HAZ was then estimated using the well-established relationships between hardness and tensile strength. It was further assumed that the plastic behavior of this material, including yielding, experienced the same ratio of reduction as the tensile strength. The reduced strength curve for the SPV490 HAZ material is plotted in Fig. 15 and was used in the posttest analysis of the local area around the large tear.

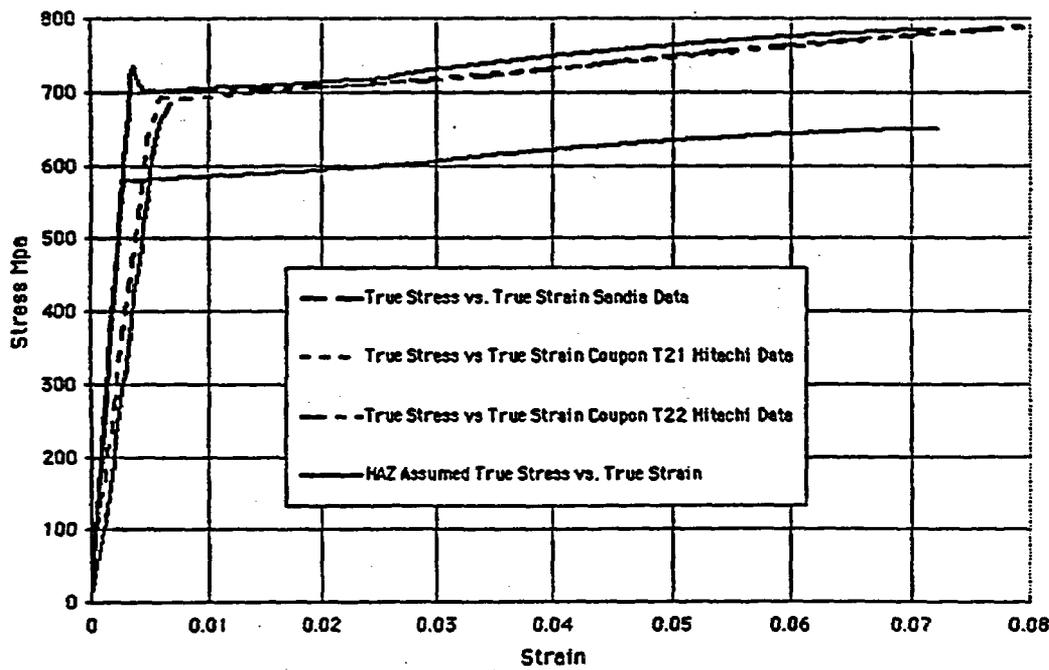


Figure 15. Stress-strain curves for the 9 mm SPV490 steel

Global Analysis Results

A global 3D finite element model with revised material models for the two steels was used to analyze the global response of the SCV model. A nominal gap size of 22 mm was used in the posttest analysis instead of the pretest gap size of 18 mm to provide a better representation of the as-built gap dimension between the SCV model and the CS.

Figure 16 shows the hoop strains as a function of model elevations at the free-field azimuth of 270° at 4.5 MPa. The posttest analysis results provided a better correlation with the measured strains than the pretest predictions. The free-field response of the SCV model (at Round Robin standard output location # 24) together with the pretest prediction and the posttest analysis results are plotted in Fig. 17. The effect of using the revised material models and the increased gap size was demonstrated by that the strain results

of the posttest analysis merged with the measured strains at a strain of about 1.7%. However, the discrepancy in simulating the onset of yielding of the model still exists.

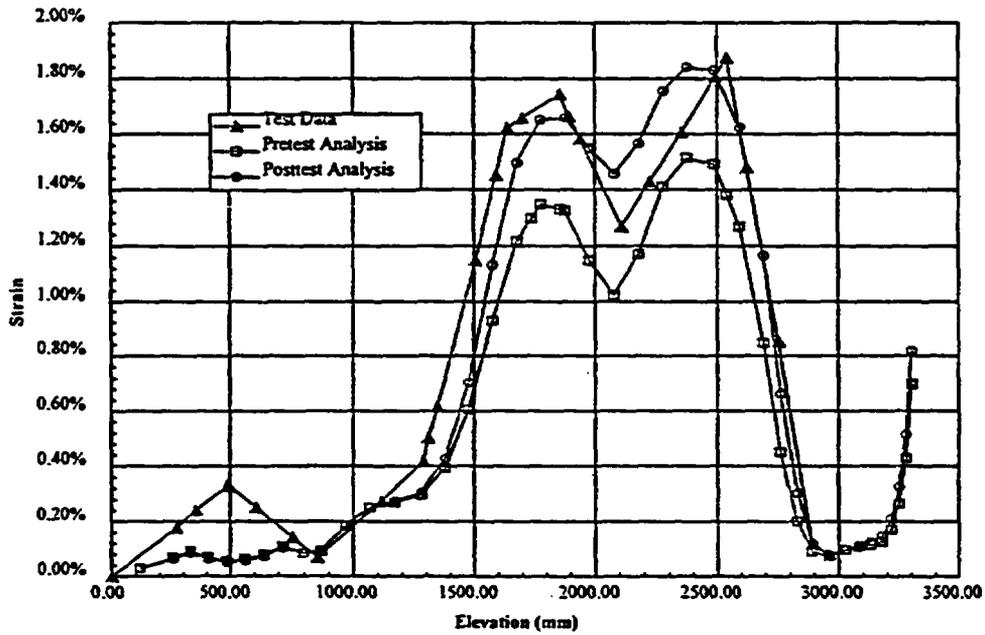


Figure 16. Hoop strain versus model elevation at 270° azimuth at 4.5 MPa

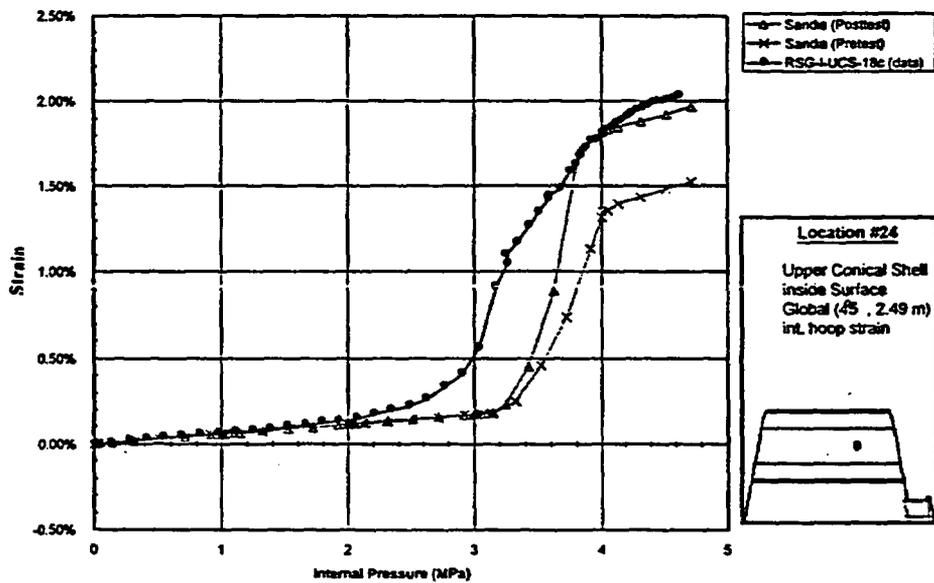


Figure 17. Interior hoop strains at upper conical shell section (@ Round Robin standard output location # 24)

Local Analysis Results Around the Equipment Hatch

The local 3D equipment hatch model was re-analyzed by assigning the elements highlighted in black in Fig. 18 with the reduced strength material model for the SPV490 HAZ. The large tear is located inside this highlighted strip of elements. The posttest analysis results, shown in Fig. 19, indicate that the highest strains appear around the large tear in the SPV490 HAZ steel shell.

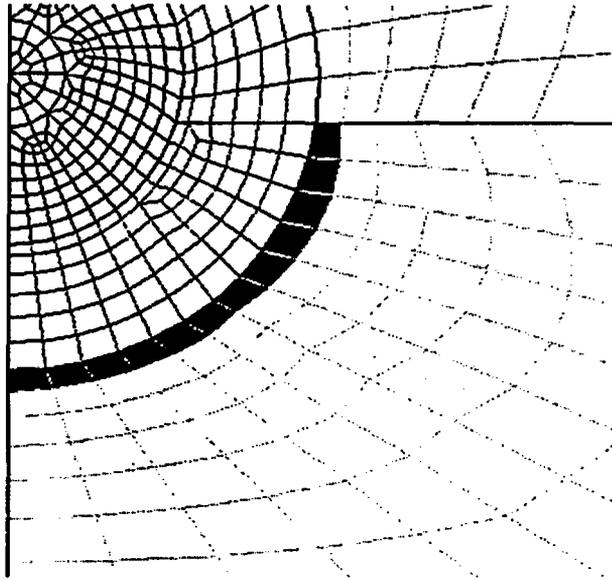


Figure 18. Local 3D equipment hatch model with SPV490 HAZ elements highlighted in black

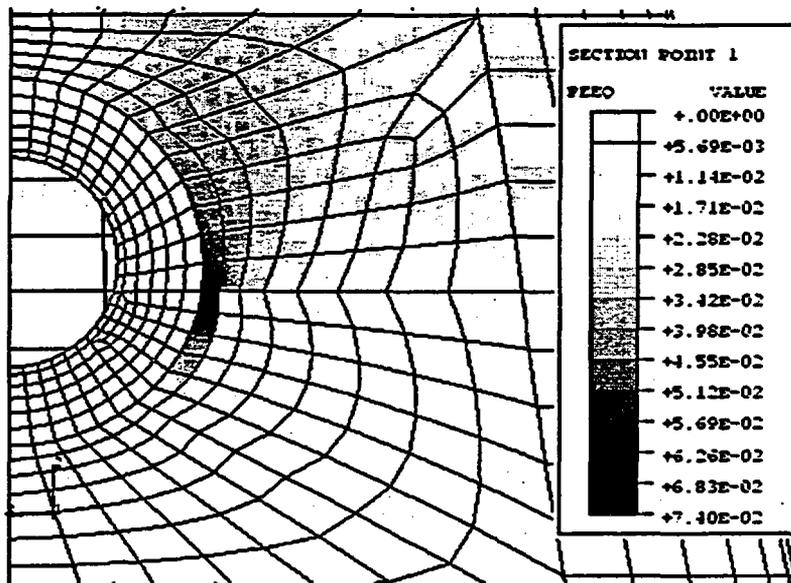


Figure 19. Equivalent plastic strain contours around the equipment hatch from the posttest analysis

Small Tear Analysis

The small tear occurred inside the weld relief opening at the middle stiffening ring (Fig. 4). The area around this tear was simulated in a local finite element model. The vertical weld seam was not modeled, and therefore there was no hardened or thickened area at the vertical centerline inside the opening. A contour plot of the equivalent plastic strains on the interior surface of the SCV model, generated by this local model, is shown in Fig. 20. The peak strains are concentrated in two areas on either side of the vertical centerline of the opening where the vertical weld seam would be located. The area of high strains coincides well with the location of the small tear.

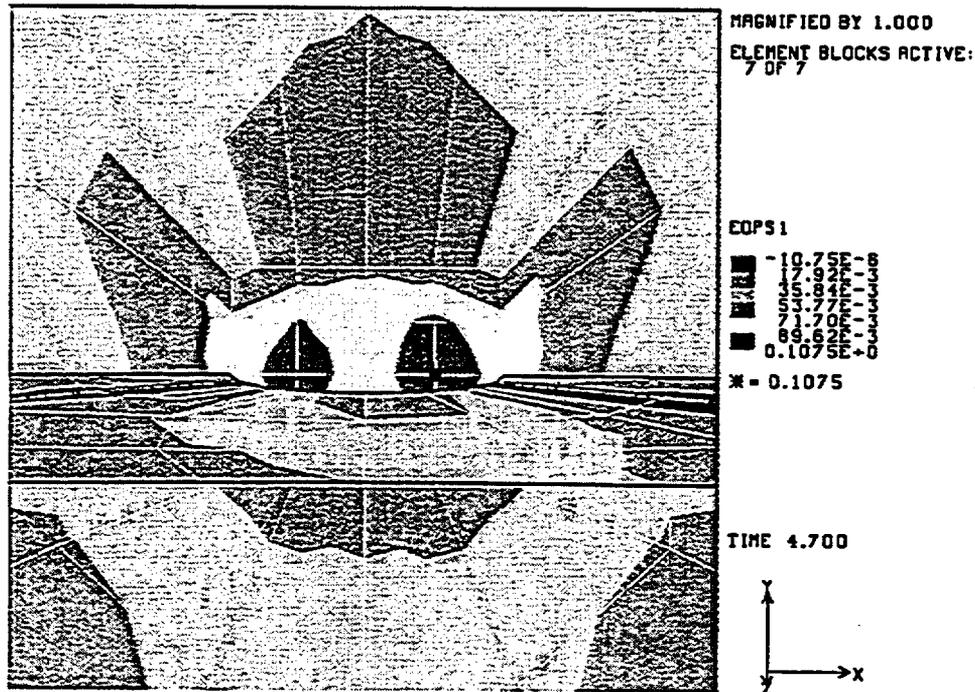


Figure 20. Contours of equivalent plastic strains on the interior surface of the SCV model adjacent to the weld relief opening at the middle stiffening ring at a pressure of 4.7 MPa

CONCLUSION

The high pressure test of the SCV model conducted at Sandia National Laboratories on December 11-12, 1996, was considered a success with regard to the specified test objectives:

1. The test provided experimental data for checking the capabilities of analytical methods well into the inelastic range of the model. While it appears that some generalized contact was occurring at the time of the failure, it is not clear that the data is adequate to confirm the validity of contact algorithms in the analysis codes.
2. The test confirmed the critical nature of discontinuities, such as penetrations, as potential failure mechanisms. The test also identified the potential significance of local changes in material properties due to welding and local fabrication details on potential failure modes. The measured

global strains at failure of 2 % are also consistent with previous tests of steel containment vessel models (Horschel, et al., 1993).

3. The test and analytical results should provide useful information for the evaluation of prototypical containment structures and focus attention on critical details and analysis methodologies.

REFERENCES

Horschel, D. S., Ludwigsen, J. S., Parks, M. B., Lambert, L. D., Dameron, R. A., and Rashid, Y. R., June 1993, "Insights Into The Behavior of Nuclear Power Plant Containments During Severe Accidents," SAND90-0119, NPRW-CON90-1, Sandia National Laboratories, Albuquerque, NM.

Luk, V. K., and Klamerus, E. W., 1996, "Round Robin Pretest Analyses of a Steel Containment Vessel Model and Contact Structure Assembly Subject to Static Internal Pressurization," NUREG/CR-6517, SAND96-2899, Sandia National Laboratories, Albuquerque, NM.

Luk, V. K., Hessheimer, M. F., Matsumoto, T., Komine, K., and Costello, J. F., August 17-22, 1997, "Testing of a Steel Containment Vessel Model", Proceedings of the 14th International Conference on Structural Mechanics in Reactor Technology, Lyon, France.

Luk, V. K., Hessheimer, M. F., Rightley, G. S., Lambert, D. L., and Klamerus, E. W., 1998 (to be published), "Design, Instrumentation and Testing of a Steel Containment Vessel Model," Sandia National Laboratories, Albuquerque, NM.

Porter, V. L., Carter, P. A., and Key, S. W., 1996, "Pretest Analyses of the Steel Containment Vessel Mode," NUREG/CR-6516, SAND96-2877, Sandia National Laboratories, Albuquerque, NM.

Van Den Avyle, J. A., and Eckelmeyer, K. H., 1998 (to be published), "Posttest Metallurgical Evaluation Results for the SCV High Pressure Test," Sandia National Laboratories, Albuquerque, NM.

NEW SEISMIC DESIGN SPECTRA FOR NUCLEAR POWER PLANTS

Robin K. McGuire¹, Walter J. Silva², and Roger Kenneally³

¹Risk Engineering, Inc.

²Pacific Engineering and Analysis

³US Nuclear Regulatory Commission

ABSTRACT

Under a US Nuclear Regulatory Commission-sponsored project recommendations for seismic design ground motions for nuclear facilities are being developed. These recommendations will take several forms. Spectral shapes will be developed empirically and augmented as necessary by analytical models. Alternative methods of scaling the recommended shapes will be included which use a procedure that integrates over fragility curves to obtain approximately consistent risk at all sites. Site-specific soil effects will be taken into account by recommending site-specific analyses that can be used to modify rock hazard curves at a site. Also, a database of strong motion records will be archived for the project, along with recommendations on the development of artificial motions. This will aid the generation of motions for detailed soil- and structural-response studies.

INTRODUCTION

In 1996 the Nuclear Regulatory Commission (NRC) amended its regulations to update the criteria used in decisions regarding nuclear power plant siting, including geologic, seismic, and earthquake engineering considerations for future applications; USNRC (1996). As a follow-on to the revised siting regulations, it is necessary to develop state-of-the-art recommendations on the design ground motions commensurate with seismological knowledge and engineering needs. The current design spectra in Regulatory Guide 1.60 (USNRC 1973) were based on limited, principally western United States earthquake strong-motion records, available at that time. Since 1996 the NRC has funded a project to develop up-to-date seismic design spectra for the US. The work combines empirical and analytical approaches, supplementing data where they are sparse using theoretical methods to develop the recommended spectra for a range of earthquake magnitudes and distances. Soil conditions necessarily involve site-specific parameters, and we demonstrate and recommend procedures to account for local soil effects on earthquake motions. A Review Panel consisting of Carl Stepp (Chair), David Boore, Allin Cornell, I.M. Idriss, and Robert P. Kennedy review the work and offer guidance on procedures. The prime contractor is Risk Engineering, Inc., with Pacific Engineering and Analysis developing databases, spectral shapes, site response procedures, and spectral matching criteria. This paper reviews the scope of the work, indicates the direction that recommendations are taking, and presents preliminary results. Final results will be available in the project report.

SPECTRAL SHAPES

Databases for the western US are available in the form of strong motion accelerograms for moment magnitudes M in the range 5.0 to 7.6 and source-to-site distances R of 1 to 200 km. Rock conditions in California are generally soft, with near-surface shear wave velocities of 200-450 m/s (700-1500 ft/sec).

The databases of strong motion records and empirical attenuation relations form the basis for recommended spectral shapes on rock for defined M and R bins, augmented as necessary by analytically derived shapes. In the application of these spectral shapes for design, the M - R combination is defined by the dominant earthquake as determined from a probabilistic seismic hazard analysis (PSHA). Examples of procedures for defining the dominant earthquake are described in McGuire (1995) and USNRC (1997).

A summary of rock and soil records from the western US is shown in Table 1, in terms of M and R bins. Also shown are preliminary summary statistics for mean peak ground acceleration (PGA), mean peak ground velocity (PGV), and mean peak ground displacement (PGD). This summary indicates the usual trends in strong motion data, i.e. that data are abundant for moderate magnitudes at moderate source-to-site distances, but are sparse for large magnitudes and short distances and small magnitudes at long distances. The former category is more troublesome from a design perspective and requires modeling for confirmation.

Central and eastern US (CEUS) strong motion records are sparse. Thus it is necessary to augment the CEUS empirical motions with analytically derived spectral shapes. This analysis uses a point- and finite-source representation of the earthquake rupture, attenuates both body and surface waves, accounts for near-surface attenuation of high frequencies, and assumes that ground motion is a band-limited, white noise process. Calibration of the model with available records confirms the underlying assumptions and provides estimates of the model parameters. One outstanding issue, however, is whether the seismic energy at the source has a "single-corner" or "double-corner" spectrum; this is the focus of independent research, and the current project will include each model as an alternative. Rock conditions in the CEUS are generally hard, with near-surface shear wave velocities generally exceeding 3000 m/s (10000 ft/sec).

Figure 1 indicates the difference in spectral shapes between the single- and double-corner models, for both the WUS and CEUS. The shapes are presented as ratios of spectral acceleration divided by PGA. CEUS shapes typically have more high-frequency content but lower SA at intermediate periods, when normalized by PGA. The double-corner model has the largest influence for CEUS spectral shapes at periods longer than 0.5s.

TABLE 1
Characteristics of WUS Records in M-R Bins (Preliminary)

| Site | M | R, km | # of spectra | mean PGA, g | mean PGV, cm/s | mean PGD, cm |
|------|-------|--------|--------------|-------------|----------------|--------------|
| Rock | 5-5.9 | 0-10 | 30 | 0.18 | 8.14 | 0.80 |
| | 6-6.9 | 0-10 | 32 | 0.44 | 32.7 | 6.22 |
| | 7+ | 0-10 | 6 | 0.93 | 81.7 | 47.4 |
| Soil | 5-5.9 | 0-10 | 24 | 0.26 | 18.6 | 3.11 |
| | 6-6.9 | 0-10 | 77 | 0.38 | 46.9 | 14.8 |
| | 7+ | 0-10 | 4 | 0.40 | 44.5 | 21.3 |
| Rock | 5-5.9 | 10-50 | 180 | 0.11 | 5.08 | 0.54 |
| | 6-6.9 | 10-50 | 238 | 0.13 | 8.81 | 1.96 |
| | 7+ | 10-50 | 6 | 0.17 | 8.80 | 2.50 |
| Soil | 5-5.9 | 10-50 | 378 | 0.11 | 6.63 | 0.87 |
| | 6-6.9 | 10-50 | 542 | 0.14 | 10.8 | 2.25 |
| | 7+ | 10-50 | 56 | 0.16 | 22.4 | 10.5 |
| Rock | 5-5.9 | 50-100 | 32 | 0.05 | 2.22 | 0.21 |
| | 6-6.9 | 50-100 | 102 | 0.06 | 3.87 | 0.79 |
| | 7+ | 50-100 | 10 | 0.06 | 5.16 | 2.64 |
| Soil | 5-5.9 | 50-100 | 42 | 0.06 | 3.11 | 0.38 |
| | 6-6.9 | 50-100 | 158 | 0.07 | 6.23 | 1.26 |
| | 7+ | 50-100 | 14 | 0.10 | 11.2 | 5.42 |

CHOICE OF SPECTRAL LEVEL

In addition to the spectral *shape*, the overall level of the spectrum must be specified. This choice may be made from a PSHA by defining a target annual frequency of exceedence for the spectrum. Alternatively the level could be defined using an acceptable annual frequency of failure P_f at the component level and convoluting the seismic hazard results with component fragility curves to relate component performance to seismic hazard. The failure frequency P_f can be represented as:

$$P_F = \int_0^{\infty} H(a) \frac{dP_{Fa}}{da} da \quad (1)$$

where $H(a)$ is the hazard curve and P_{Fa} is the probability of failure (the "fragility") given ground motion amplitude "a", which captures both response and capacity uncertainties.

With some realistic assumptions on the shape of the hazard curve and the fragility curve, it is possible to derive a simple expression for P_f . First we assume that the hazard curve $H(a)$ is linear on log-log scale, i.e.

$$H(a) = ka^{-K_H} \quad (2)$$

where a is spectral acceleration level, k is a constant, and K_H is the slope of the hazard curve in log-log space. Actual hazard curves tend to get steeper at higher amplitudes, but over the important range of amplitudes for P_f calculations they can be approximated as linear on log-log scale.

Second we assume that component fragilities are lognormally distributed. This means that

$$P_{Fa} = \int_0^a \frac{1}{y\sqrt{2\pi}\beta} \exp\left\{-\frac{(\ln y - \overline{\ln y})^2}{2\beta^2}\right\} dy \quad (3)$$

where $\overline{\ln y} = \ln CAP_{50}$ the median component capacity, and β is the logarithmic standard deviation of capacity.

Substituting equations (2) and (3) into (1) gives

$$P_F = \int_0^{\infty} k a^{-K_H} \frac{1}{a\sqrt{2\pi\beta}} \exp\left\{-\frac{(\ln y - \overline{\ln y})^2}{2\beta^2}\right\} da \quad (4)$$

Transforming the integration variable a to variable $x = \ln a$ gives

$$P_F = \frac{k}{\sqrt{2\pi\beta}} \int_{-\infty}^{\infty} \exp\{-K_H x\} \exp\left\{-\frac{(x - \overline{\ln y})^2}{2\beta^2}\right\} dx \quad (5)$$

The integrand above is in the form

$$\exp\{cx\} Z(x) \quad (6)$$

where c is a constant and $Z(x)$ is the normal density function. The definite integral equation (5) can be solved by expansion or by published methods of integrating functions of normal probability distribution (e.g. Owen, 1980), yielding

$$P_F = k C A P_{50}^{-K_H} \exp\left\{\frac{1}{2}(K_H \beta)^2\right\} \quad (7)$$

This form, designated the "risk equation," was first derived by G. Toro and published in Sewell et al. (1991, 1996). Expressing the hazard H_s at a ground motion level a^* corresponding to a safe-shutdown ground motion (SSGM), using equation (2) gives:

$$H_s = k(a^*)^{-K_H} \quad (8)$$

Solving for k and substituting into equation (7) gives:

$$P_F = H_s (a^*)^{K_H} C A P_{50}^{-K_H} \exp\left\{\frac{1}{2}(K_H \beta)^2\right\} \quad (9)$$

We can now derive a probability ratio R_p , as the probability H_i that a^* will be exceeded, divided by the probability of failure P_F :

$$R_p = H_i / P_F \quad (10)$$

This ratio is usually much greater than unity because P_F is much less than the hazard at a^* . R_p can be expressed as:

$$R_p = \left(\frac{CAP_{50}}{a^*} \right)^{K_H} \exp \left\{ -\frac{1}{2} (K_H \beta)^2 \right\} \quad (11)$$

Instead of using the median capacity CAP_{50} to designate capacity, we can use the "high confidence of low probability of failure" value, or HCLPF, where for a lognormal distribution the two are related by

$$HCLPF = CAP_{50} \exp \{ -x_p \beta \} \quad (12)$$

where x_p is the number of standard deviates corresponding to the frequency of failure at the HCLPF, which is 2.326 for 1% frequency of failure. Also, we can express the required HCLPF in terms of a^* times a factor of safety F_R :

$$HCLPF = SSGM \cdot F_R \quad (13)$$

Solving these last two equations for CAP_{50} and SSGM, and substituting into equation (11) gives:

$$R_p = F_R^{K_H} \exp \left\{ x_p K_H \beta - \frac{1}{2} (K_H \beta)^2 \right\} \quad (14)$$

This gives a simple means to calculate P_F , given that the hazard associated with the SSGM is known. The probability ratio R_p depends on the factor of safety F_R , the hazard curve slope K_H , and β of the fragility function; for the HCLPF defined at the 1% frequency of failure point, $x_p = 2.326$ as explained above.

Equation (14) also gives an easy way to compute the effect of hazard curve slope and fragility β on P_F for a specified hazard corresponding to a selected UHS. Stated another way, if we pick a UHS at each site with the same annual probability of exceedence, and define the HCLPF in terms of equation (13), equation (14) allows us to examine the *risk consistency* across sites for different hazard curve slopes K_H and fragility uncertainties β . The use of equation (14) in this way is demonstrated below.

A couple of points about the *distributions* of $H(a^*)$ and P_F are important. $H(a^*)$ is uncertain because of lack of knowledge in the earth sciences about earthquake sources, ground motions, etc. This uncertainty has been quantified by EPRI and LLNL at EUS plant sites and by utilities at several WUS plant sites. If we use the mean of this distribution we will achieve a mean P_F for any set of design rules. The mean has the advantage that we can compute (and control) the mean P_F for multiple plants. That is, we have n plants and a total acceptable probability of component failure at these plants, we can achieve that by specifying a mean P_F at each plant. The disadvantage is that the mean is sensitive to low probability, high consequence assumptions in the seismic hazard analysis and is not as stable (from study to study) as the median.

If we use the median $H(a^*)$ we will achieve an approximate median P_F . The median has the advantage that it is more stable than the mean, but a target mean or median P_F over n plants cannot readily be translated to a required median P_F at each plant. So use of the median $H(a^*)$ leads to ill-constrained limits on P_F over multiple plants. For this reason the use of the mean $H(a)$ curve is recommended.

A final point is that R_p can be controlled by "deterministic acceptance criteria" associated with design codes and guides, and by a "scale factor" that moves the capacity up or down as a function of the hazard curve slope K_H , the desired P_F , or the desired R_p for a given $H(a)$. This scale factor is conveniently thought of as a scaling of the UHS to specify an SSGM spectrum. The total factor of safety F_R , is then α times SF, where α is the conservatism achieved by design procedures (e.g. 1.67 on the HCLPF) and SF is the scale factor. The SSGM is then the UHS scaled by SF. It is appropriate to define SF to scale the UHS to account for the site-specific (and natural period-specific) slope of the hazard curve. R.P. Kennedy (personal communication, 1997) has suggested the following scale factor:

$$SF = \max\{0.7, 0.35A_R^{1.2}\} \quad (15)$$

where $A_R = [\log K_H]^{-1}$. Thus A_R increases as the hazard curves become more shallow, so SF increases, i.e. the design values become higher for shallow hazard curves. With this definition, the SSGM can be thought of as:

$$SSGM = UHS \times SF \quad (16)$$

i.e., the SSGM is the UHS "corrected" for the slope of the hazard curve. For $A_R = 2.40$ (which corresponds to slope $K_H = 2.63$), $SF = 1$, i.e. the SSGM equals the UHS.

Another way to look at the design is through the total factor of safety F_R (see equation (15)). If the amount of conservatism in design codes and guides (sometimes referred to as the "deterministic acceptance criterion") is 1.67, then the total factor of safety F_R is:

$$F_R = 1.67SF \quad (17)$$

The advantage of using a slope-dependent scale factor SF as defined in equation (15) is demonstrated in the next section.

RESULTS FOR EXAMPLE SITES

To test several methods for risk-consistent spectra, we examined eleven sites and three ground motion measures at each site, shown in Table 2:

TABLE 2
Sites and Ground Motion Measures Used for Testing Procedures

| <u>No.</u> | <u>Site</u> | <u>Measure</u> | <u>No.</u> | <u>Site</u> | <u>Measure</u> |
|------------|----------------|----------------|------------|----------------|----------------|
| 1 | Arkansas plant | PGA | 17 | Shearon Harris | SV 1 Hz |
| 2 | Arkansas plant | SV 1Hz | 18 | Shearon Harris | SV 10 Hz |
| 3 | Arkansas plant | SV 10 Hz | 19 | Susquehanna | PGA |
| 4 | Browns Ferry | PGA | 20 | Susquehanna | SV 1 Hz |
| 5 | Browns Ferry | SV 1 Hz | 21 | Susquehanna | SV 10 Hz |
| 6 | Browns Ferry | SV 10 Hz | 22 | Vogtle | PGA |
| 7 | Davis Besse | PGA | 23 | Vogtle | SV 1 Hz |
| 8 | Davis Besse | SV 1 Hz | 24 | Vogtle | SV 10 Hz |
| 9 | Davis Besse | SV 10 Hz | 25 | Zion | PGA |
| 10 | Maine Yankee | PGA | 26 | Zion | SV 1 Hz |
| 11 | Maine Yankee | SV 1 Hz | 27 | Zion | SV 10 Hz |
| 12 | Maine Yankee | SV 10 Hz | 28 | California | PGA |
| 13 | Seabrook | PGA | 29 | California | SV 1 Hz |
| 14 | Seabrook | SV 1 Hz | 30 | California | SV 10 Hz |
| 15 | Seabrook | SV 10 Hz | 31 | Washington | PGA |
| 16 | Shearon Harris | PGA | 32 | Washington | SV 1 Hz |
| | | | 33 | Washington | SV 10 Hz |

For the first 27 sets of results we used the LLNL hazard curves calculated for the USNRC (Sobel, 1994). For the "California" site, we calculated hazard at a site located near Santa Maria, California (120.5° W, 35.0° N), which has high frequencies dominated by nearby faults and long periods dominated by the more distant San Andreas fault. (A repeat of the 1857 earthquake dominates the long period hazard at this site.) For ground motion estimation the attenuation equation of Abrahamson and Silva (1995) was selected.

The last site examined was in Washington, located at 121°W and 46°N. This is in south-central Washington and also has high frequencies dominated by local earthquakes and low frequencies dominated by a large earthquake. In this case a large subduction zone earthquake controls the long-period hazard. We model this event using the assumptions of the US Geological Survey for the national seismic hazard

maps. That is, an earthquake of $M=9$ occurs in the subduction zone with rate 1/500 per year (credibility 1/3), or earthquakes of $M=8$ to 9 occur with rate 1/110 per year (credibility 2/3). For both the California and Washington sites we model local earthquakes with the US Geological Survey gridded seismicity, as well as local faults for the California site.

Calculations were made of the probability ratio R_p for the 33 site-parameter combinations listed above. This is an appropriate parameter to use because, if we start with the same hazard level $H(a^*)$ at all sites and all natural periods, and achieve a consistent R_p with our procedure, we will achieve a consistent probability of failure P_F .

Figure 2 shows R_p values for the 33 site-parameter combinations, calculated using the *mean* hazard curve for each site. For this plot R_p was calculated from equation (eq. 14) using the derivation from the risk equation. The top plot in Figure 2 shows R_p when the SSGM is taken to be *equal* to the UHS at the natural period of the parameter; the bottom plot shows R_p when the SSGM = UHS x SF, as in equation (16). The scale factor SF really helps the consistency across sites and across parameters; results without SF vary from about 6 to 48 (a factor of 8), but with SF they vary from about 18 to 45 (a factor of 2.5). This remaining factor of 2.5 is the effect of β . It would be inappropriate to define the SSGM on the basis of component response, since that would require multiple design spectra for a single facility.

PROCEDURES FOR SOIL SITES

Results presented above assume that facility design is for a rock site. If soil conditions exist at a site, modifications will be necessary to derive the appropriate design level. Several options are being considered; these follow ideas expressed by Cornell (1996) and Cornell and Bazzurro (personal communication, 1997).

Option 1: Direct approach.

This approach models soil response directly as a function of M and R (through a site specific attenuation relation), to calculate soil hazard curves $H_s(a)$ as

$$H_s(a) = \int P_s[A > a | M, R] f_{M,R} dm dr \quad (18)$$

This has the advantage of directness and consistency with the derivation of rock hazard curves. The disadvantages, however, make this approach unworkable for most sites. First, the seismic hazard analysis cannot be conducted prior to obtaining detailed soil-specific information (shear wave velocities, modulus and damping curves) for each location where facilities are to be designed at the site. This procedure couples the design criteria process with the collection of site-specific information, and preliminary designs based on approximate amplification factors would be awkward. Second, if preliminary site information were obtained and later refined, the entire seismic hazard analysis must be repeated and documented to incorporate the new information. Finally, this option has not been used in the past, at least for site-specific soil properties, and implementing it would require addressing issues of consistency and accuracy in representing soil response with a generic attenuation equation form. These disadvantages are not insurmountable, but they imply that a different approach would be more efficient.

Option 2: Simple scale factor

In this approach, the hazard curve $H(a)$ for rock is represented as spectral acceleration $a_R(h)$ and is simply scaled up (or down) at each frequency by a soil-dependent shape:

$$a_s(h) = a_r(h) \overline{S(a_R, f)} \quad (19)$$

where $\overline{S(a_R, f)}$ is a mean scale factor that depends on a_R and frequency f , and is developed through site specific response analyses. $\overline{S(a_R, f)}$ can accommodate variabilities in site amplification caused by uncertainties in soil properties, for example. This approach has the advantage of simplicity, but the disadvantage of inaccuracy. $S(a_R, f)$ is a function not only of a_R and f but also of M (but probably not of distance R), because soil amplification depends on characteristics of the ground motion such as duration and frequency content (not just at frequency f). The dominant contribution by M changes as a function of a_R level, source contribution to hazard, and other factors.

Option 3: More detailed scale factor.

The third option is to develop a more detailed scale factor that accounts for additional features of ground shaking, but allows the simplicity of using hazard curves developed on rock:

$$P[A_s > a_s] = \iint f_{a_R}(a_R, M, \dots) P[S > \frac{a_s}{a_r} | a_r, f, M, \dots] da_r, \dots \quad (20)$$

where $f_{a_R}(a_R, M, \dots)$ is the rock hazard curve in density form, and the integral is over all factors M, \dots that are used to develop the scale factor. The hazard results $f_{a_R}(a_R, M, \dots)$ can be obtained from a standard seismic hazard analysis where results have been deaggregated by the contributions by magnitude, distance, etc.

The advantage of equation 20 is that additional uncertainties, most importantly on soil characteristics and their effects on amplification, can be incorporated into $P[S > \frac{a_s}{a_r} | a_r, f, M, \dots]$.

The project is currently examining alternative definitions of scale factors under this option and their accuracies by comparing with Option 1 (the direct approach) for several sites and sets of soil properties. Recommendations will be made based on the most workable option that can be implemented to give accurate estimates of soil hazard.

CONCLUSIONS

To calibrate design spectral shapes, strong motion records are available for the WUS over a range of magnitudes and distances. There are still comparatively few records for large magnitudes and short source-to-site distances, however. For the CEUS, all magnitudes and distances lack sufficient empirical strong motion data, and these records will be generated with ground motion models calibrated to replicate available ground motion characteristics at smaller magnitudes and longer distances.

An outstanding issue is the use of the single-corner or double-corner model of the seismic source spectrum. Results will be presented in this project for both models, anticipating that several years will be required to achieve resolution of this issue.

A method has been developed and tested to determine the *amplitude* of ground motion for design, as well as the shape. Results using test sites in the CEUS and on the west coast, and using typical component fragility characteristics, indicate that the annual frequency of component failure is about 15 to 45 times less than the annual frequency of exceedence of the design spectrum, using realistic design procedures. This means that, for example if the median frequency of exceedence of a site's design spectrum is 1×10^{-4} , the median component frequency of failure is about 3×10^{-6} . The ultimate choice of a recommended spectral level must be made with a combination of analysis to determine acceptable failure frequencies, calibration to accepted existing design procedures, and judgment.

Recommending spectral shapes for soil sites requires additional procedures. One straightforward method is to conduct the PSHA with site-specific soil attenuation equations, to obtain seismic hazard curves and uniform hazard spectra (UHS) for the soil surface. However, as recommended in NRC (1997), it is often more practical to conduct the PSHA for rock outcrop conditions and later translate these to soil surface motions, because various facilities may be located on different soils, or detailed site-specific data may not be available early in the project. In this case a site's rock UHS at a target annual frequency of exceedence can be translated to a soil UHS at the same or similar frequency of exceedence, accounting for uncertainties in the soil properties. Procedures to accomplish this will be demonstrated in the project.

In addition to recommended spectral shapes, the project will archive a database of strong motion records for the recommended M and R bins, for both rock and soil conditions. These will be empirical records for bins where data are abundant, augmented by artificial motions derived to have the correct frequency content for bins where data are sparse or non-existent.

A final set of recommendations concerns criteria to match artificial motions to recommended spectral shapes and levels. Such motions might be used for input to detailed dynamic analyses of building response, for example. The recommended procedures for developing artificial motions concentrate on matching response spectral amplitudes at multiple frequencies and dampings, and put less emphasis on matching power spectral density functions.

This NRC-sponsored project will offer a number of recommendations on choosing spectral shapes, selecting design levels, and generating time histories of motion for the design of nuclear facilities. The objective is to achieve consistent design levels across the country for a range of seismic environments and site conditions. Procedures developed in this project to define ground motion for a risk-consistent, performance-based design are an integral part of the recommendations. A second objective is to make the

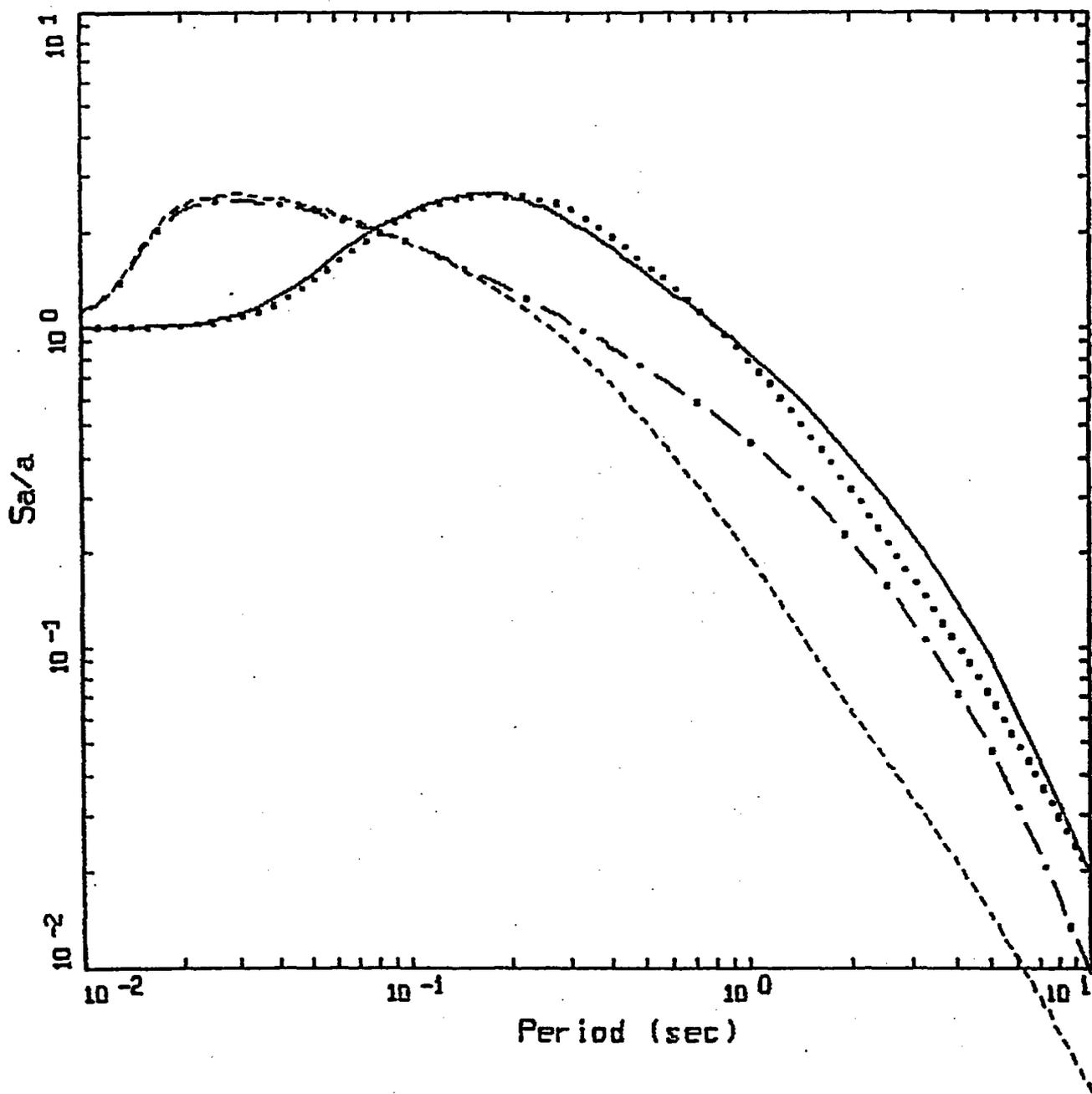
procedures easy-to-understand and technically justified, so that they will be readily accepted. There is a need to strike a balance between the engineering conservatism required to achieve the safe design, seismological knowledge, and preservation of important earthquake ground motion characteristics, such that realistic responses are considered. The results from this research will also provide tools for the seismic design of non-reactor facilities and will influence the design of non-nuclear facilities.

REFERENCES

- Cornell, C.A. (1996). "Calculating Building Seismic Performance Reliability: A basis for multi-level design norms," Paper 2122, 11th World Conf. On Earthquake Eng., Acapulco.
- McGuire, R.K. (1995). "Probabilistic seismic hazard analysis and design earthquakes: closing the loop," *Bull. Seism. Soc. Am.*, 85, 5, 1276-1284.
- Owen, D.B. (1980). "A Table of Normal Integrals," *Commun. Statist. - Simula. Computa.*, B9(4), 389-419.
- Sewell, R.T., G.R. Toro and R.K. McGuire (1991). "Impact of Ground Motion Characterization on Conservatism and Variability in Seismic Risk Estimates," Risk Eng., Inc., Report to USNRC, May.
- Sewell, R.T., G.R. Toro and R.K. McGuire (1996). "Impact of Ground Motion Characterization on Conservatism and Variability in Seismic Risk Estimates," US Nuc. Reg. Comm., Report NUREG/CR-6467, July.
- Sobel, P. (1994). *Revised Livermore seismic hazard estimates for sixty-nine nuclear power plant sites east of the Rocky Mountains*, U.S. Nuclear Regulatory Commission, Rept NUREG-1488, April.
- USNRC (1993). "Design Response Spectra for Seismic Design of Nuclear Power Plants," Regulatory Guide 1.60, Revision 1, December.
- USNRC (1996). "Reactor Site Criteria including Seismic and Earthquake Engineering Criteria for Nuclear Power Plants," *Federal Register*, Vol. 61, p 65157, December 11.
- USNRC (1997). "Identification and characterization of seismic sources and determination of safe shutdown earthquake ground motion," USNRC, Regulatory Guide 1.165, March.

DISCLAIMER

This paper was prepared in part by an employee of the United States Nuclear Regulatory Commission. It presents information that does not currently represent an agreed-upon staff position. NRC has neither approved nor disapproved its technical content.

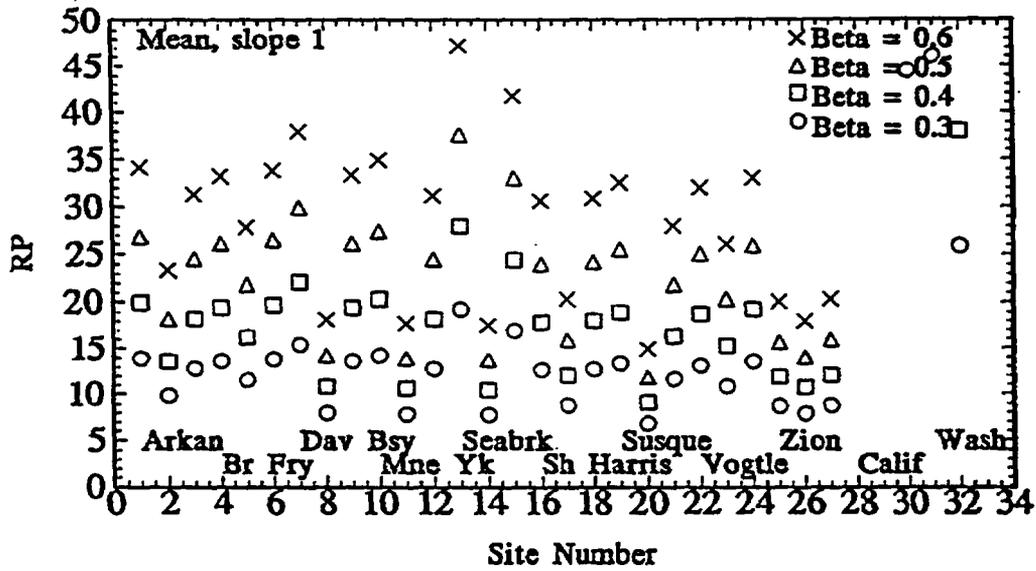


LEGEND

- WNA, 1-corner source model
- WNA, 2-corner source model
- CEUS, 1-corner source model
- CEUS, 2-corner source model

Figure 1: Normalized rock spectral shapes for WUS and CEUS 1-corner and 2-corner models, $M = 6.5$ and $R = 25$ km

HCLPF1 design using factor 1.67, ALL



HCLPF2 design using factor 1.67 x SF, ALL

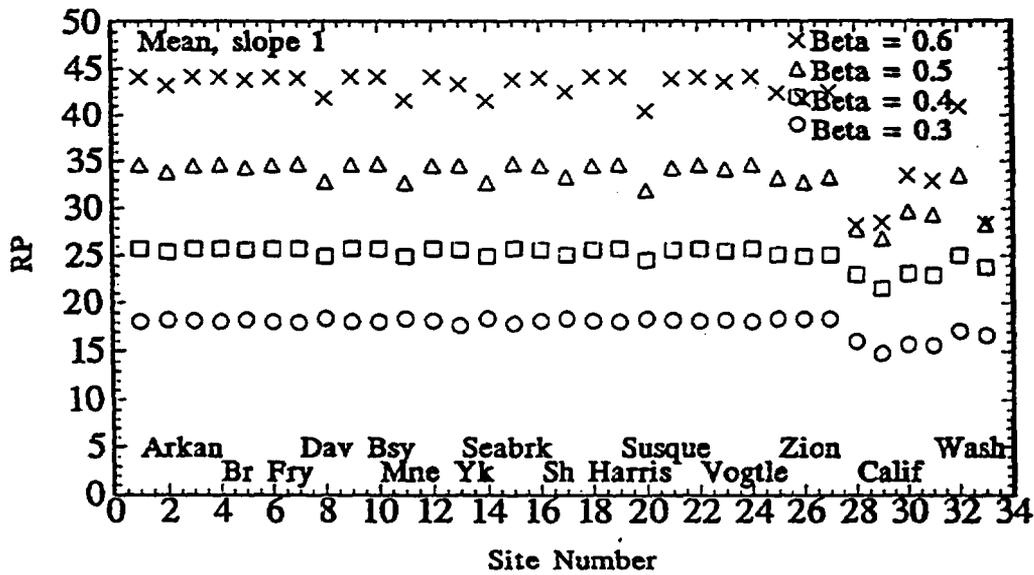


Figure 2: Ratio $R_p = H_s / P_F$ for eleven sites and three ground motion measures. Top: without slope-dependent scale factor. Bottom: with slope-dependent scale factor.

Damage mechanics-based assessment of time-dependent structural deterioration and reliability

Bruce R. Ellingwood

Department of Civil Engineering, Johns Hopkins University, Baltimore, MD 21218

Baidurya Bhattacharya

Advanced Analysis Department, American Bureau of Shipping, Houston, TX 77060

Abstract

Steel containments and liners in nuclear power plants (NPPs) may be exposed to aggressive service and environmental effects over a 40-year service life, and may be subject to corrosion, elevated temperature creep, low-cycle fatigue, and load-induced inelastic deformation. While corrosion is reasonably well-understood, the same cannot be said about the mechanisms underlying the other three processes. The initial stages of these processes often occur without perceptible manifestation, so that a significant fraction of the service life or margin of safety may already be exhausted before damage detection. Many methods for modeling structural deterioration require a measurable flaw to be applicable, and most are empirical or semi-empirical in nature. Finally, structural damage growth is an intrinsically random process. This paper explores the use of continuum damage mechanics (CDM) as a tool for evaluating damage accumulation in steel pressure boundary structures. CDM is particularly well-suited for analyzing damage that occurs over an extended period of time without visible manifestation. The governing damage growth laws are derived from the fundamental principles of thermodynamics and mechanics and are significantly less empirical in nature. This approach extends naturally into the stochastic domain, and can be integrated with time-dependent reliability assessments. The estimated conditional failure rates for structural components increase in a nonlinear fashion with time. Neglecting this nonlinear behavior may lead to an erroneous appraisal of time-dependent margins of safety.

1. INTRODUCTION

Steel containments and liners in nuclear power plants (NPPs) may be exposed to aggressive service and environmental effects over their service lives. Among the mechanisms having the potential to cause such steel pressure boundary structures to deteriorate in service are corrosion, elevated temperature creep, low-cycle fatigue, and load-induced inelastic deformation. While corrosion is reasonably well-understood, the same cannot be said about the underlying mechanisms giving rise to creep, fatigue, or inelastic deformation damage. Moreover, the initial stages of such damage often occur without perceptible manifestation. By the time that damage reaches a detectable stage, a significant fraction of the remaining service life or residual strength (or margin of safety) may already have been exhausted. Condition assessment of a containment metallic pressure boundary should provide quantitative evidence that structural performance will continue to meet or exceed a minimum standard of acceptability in the foreseeable future. Quantitative evaluation of the effects of damage accumulation on time-dependent

structural behavior is difficult. Many methods for modeling structural deterioration require a measurable flaw to be applicable. Most are empirical or semi-empirical in nature, and rely heavily on experimental data. When such data are limited or unavailable, extrapolation to service conditions for service life prediction is difficult. Finally, structural damage growth is an intrinsically random process; yet, most available approaches to random damage growth tackle the problem by simply "randomizing" the corresponding deterministic models, instead of investigating the actual sources of the randomness.

Time-dependent structural reliability analysis provides the framework for integrating information on material and structural degradation and damage accumulation, service and environmental factors and nondestructive evaluation technology. Research in progress, supported by Oak Ridge National Laboratory and the US Nuclear Regulatory Commission, is aimed at: identifying mathematical models to evaluate structural degradation and damage accumulation; recommending statistically-based sampling plans for nondestructive evaluation; and assessing the probability that structural capacity will not degrade to an unacceptable level during a future service period (Naus, et al, 1996; Ellingwood, et al, 1996; Oland and Naus, 1998).

A recent phase of this research has explored the use of the relatively new field of continuum damage mechanics (CDM) as a tool for evaluating damage accumulation in steel pressure boundary structures (Bhattacharya and Ellingwood, 1998c). CDM deals with the aggregate effects of micro-structural defects, expressed in terms of quantities that are observable at the structural level, e.g., changes in the elastic modulus or stiffness. It is particularly well-suited for analyzing damage that may occur over an extended period of time without visible manifestation. CDM can also address some of the fundamental aspects of random structural damage growth. CDM has the potential to reduce the level of empiricism associated with other approaches to modeling structural damage accumulation.

1.1 Basic Definitions in CDM

In continuum damage mechanics (CDM), damage is defined as the effective density of defects/discontinuities on a cross-section in a given orientation (Lemaitre, 1985). Damage is generally a tensor due to its directional nature (Krajcinovic, 1984); however, it is common to model damage as *isotropic*, under the assumption that the effective fractional loss of area is the same regardless of the orientation of the cross-section. Damage is considered to be isotropic in this paper and is described by a scalar, D , taking values between 0 and 1. The constitutive law for a damaged material can be derived from the concept of effective stress and the principle of strain equivalence (Lemaitre, 1985; Kachanov, 1986; Chaboche, 1988). The effective stress is defined as

$$\tilde{\sigma} = \frac{\sigma}{1-D} \quad (1)$$

where σ is the nominal stress. The strain equivalence principle asserts that the strain response of an undamaged body under the effective stress is the same as that of a comparable damaged body under the nominal stress. Applying this to uniaxial elastic deformation, the damage variable may be related to the fractional loss in stiffness:

$$D = 1 - \frac{\tilde{E}}{E} \quad (2)$$

where \tilde{E} is the elastic modulus of the damaged material, and E is the elastic modulus of a comparable undamaged material. Eq (2) allows measuring the extent of damage in a structural component by one of several conventional non-destructive methods, including direct tension tests, ultrasonic pulse velocity, measurement of electrical resistivity, etc (Lemaitre, 1992).

1.2 Critical Damage

Damage accumulation is a thermodynamically irreversible (i.e., dissipative) process, and the damage variable should be a non-decreasing function of time (assuming no corrective intervention). Failure occurs when D reaches the critical damage, D_c . In CDM, "failure" is not necessarily fracture, but is the condition when the essential assumption that damage arises out of a volume-wide degradation of the microstructure ceases to be applicable. At this point, the damage-causing process becomes localized and produces a dominant defect. Subsequent damage analysis then can be performed by other methods, e.g., fracture mechanics.

This concept of failure allows D_c to have values less than unity, unlike many phenomenological models (like Miner's rule in fatigue) in which cumulative damage is postulated as equal to 1 at failure. D_c is postulated as a fundamental material property (e.g., Chow and Wei, 1991) that may be dependent on temperature but is otherwise independent of the loading history. Hence D_c determined from one experiment (e.g., a simple tension test) for a particular material at a given temperature can be used to predict failure in a more complex loading situation (e.g., in high cycle fatigue). Experimentally determined values of D_c range anywhere between 0.15 and 0.85 for many metals (e.g., Lemaitre, 1992). In a stochastic analysis, D_c should generally be treated as a random variable.

2. THERMODYNAMIC BASIS OF RANDOM DAMAGE ACCUMULATION

For a deformable body \mathfrak{R} (defined by the closed boundary $\partial\mathfrak{R}$) in diathermal contact with a heat reservoir at constant temperature θ , subject to pre-localization damage-causing processes, the Helmholtz free energy, $\Psi(\theta, \underline{\epsilon}, D)$, is a function of the temperature, the symmetric strain tensor $\epsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})$ in which u_i is displacement, and the damage variable. The rapid and continuous transitions and interactions in the microstates of the system \mathfrak{R} (Callen, 1985; Ostoja-Starzewski, 1989), and the spatial inhomogeneity at the local scale (even in a nominally homogeneous material), suggest that the Helmholtz free energy should be described as a stochastic process (Bhattacharya and Ellingwood, 1998a):

$$\Psi(t) = \int (\dot{W} - \dot{K}_E) dt - \int \Gamma dt + \int \dot{B} dt \quad (3)$$

where W is the work done on \mathfrak{R} , K_E is its kinetic energy, and Γ is the dissipation rate. The superscript dot represents time-derivative. $B(t)$ is a stochastic process representing the random fluctuation in the free energy, and $\dot{B}(t)$ is its derivative in the mean-square sense. Spatial fluctuations in the free energy at a given instant are neglected as they are assumed small in a nominally homogeneous material undergoing isotropic damage accumulation prior to localization.

Let us assume that the initial state (at time t_0) is one of thermodynamic equilibrium, and damage accumulation, though irreversible, occurs sufficiently close to equilibrium in the pre-localization stage. Under these assumptions, the first variation in $\Psi(t)$, which is generally non-zero for a system yet to achieve equilibrium, may be assumed to vanish:

$$\delta\Psi = \delta I_1 - \delta I_2 \equiv 0 \quad (4)$$

where

$$\begin{aligned} I_1 &= \dot{W} - \dot{K}_E + \frac{\partial\Psi}{\partial D} \dot{D} + \dot{B} \\ I_2 &= \dot{W} - \dot{K}_E - \frac{\partial\Psi}{\partial \underline{\varepsilon}} \dot{\underline{\varepsilon}} \end{aligned} \quad (5)$$

The validity of Eq (4) is confirmed subsequently with experimental data.

Applying an appropriate set of variations, $\delta\dot{u}_i$, to the velocity field consistent with the boundary conditions the second integral, δI_2 , can be written as:

$$\delta I_2 = \int_{\mathfrak{R}} (F_i + \sigma_{ij,j} - \rho a_i) \delta\dot{u}_i dV + \int_{\partial\mathfrak{R}_1} (T_i - \sigma_{ij} n_j) \delta\dot{u}_i d\eta \quad (6)$$

where F_i and T_i ($i=1,2,3$) are, respectively, the body forces (on \mathfrak{R}) and surface forces (on the free surface $\partial\mathfrak{R}_1$); a_i and \dot{u}_i are, respectively, the acceleration and the velocity; ρ is the mass density; and the symmetric stress tensor $\sigma_{ij} = \partial\psi / \partial\varepsilon_{ij}$, where ψ is the Helmholtz free energy per unit volume. Both integrands in Eq (6) are equal to zero as they constitute equilibrium equations of a damaged body (Krajcinovic and Sumarac, 1987). Thus the second term in Eq (4) vanishes. Hence, the first term in Eq (4) must also vanish. It can then be shown that

$$\begin{aligned} F_i - (\psi_D D'_{ij})_{,j} - \rho a_i &= 0 \quad \text{on } \mathfrak{R} \\ T_i + \psi_D D'_{ij} n_j &= 0 \quad \text{on } \partial\mathfrak{R}_1 \end{aligned} \quad (7)$$

This set of coupled partial differential equations may be difficult to solve for a body subjected to multiaxial straining. However, under uniaxial straining, a single stochastic differential equation can be derived which is amenable to closed-form solutions for different modes of damage accumulation. Since material properties and random damage growth data are available mainly for uniaxial loading conditions, this SDE is useful for testing the validity of the approach for modeling random structural damage accumulation.

3. ISOTROPIC RANDOM DAMAGE GROWTH

Under uniaxial loading, the second part of Eqs (7) reduces to

$$\sigma_\infty + \psi_D \frac{dD}{d\varepsilon} + s_b = 0 \quad (8)$$

where σ_∞ is the far-field stress (generally random) acting normal to the surface. The term s_b has dimensions of energy per unit volume per unit strain (or units of stress), and can be interpreted as a

random fluctuation imposed on the nominal stress field existing within the deformable body. Suppose that (i) s_b is a zero-mean process that assumes positive and negative values with equal probability, (ii) the mean-square fluctuation is independent of strain (or time), and (iii) the rate of fluctuation in s_b can be described as extremely rapid in comparison with the macroscopic rate of change in damage. These are satisfied if s_b is described by the Langevin equation,

$$\frac{ds_b}{d\varepsilon} = -c_1 s_b + \sqrt{c_2} \xi(\varepsilon) \quad (9)$$

where $\xi(\varepsilon)$ is a Gaussian white noise indexed with strain so that $\xi(\varepsilon) = dW(\varepsilon)/d\varepsilon$ where $W(\varepsilon)$ is the standard Wiener process; and c_1, c_2 are positive constants. Since the scale of fluctuations in s_b are short compared to the scale of the index parameter (time or strain intervals of engineering interest), we can write the following stochastic differential equation (SDE) for damage growth (Gardiner, 1985):

$$\frac{dD}{d\varepsilon} = -\frac{\sigma_\infty}{\psi_D} + \frac{\sqrt{c_2}/c_1}{\psi_D} \xi(\varepsilon) \quad (10)$$

Alternately, random damage growth may be indexed with time, rather than with strain, if the strain rate is known:

$$\frac{dD}{dt} = -\frac{\sigma_\infty}{\psi_D} \dot{\varepsilon} + \frac{\sqrt{c_4}/c_3}{\psi_D} \dot{\varepsilon} \xi(t) \quad (11)$$

where $\xi(t)$ is a Gaussian white noise indexed with time, and c_3, c_4 are positive constants defining a Langevin equation similar to Eq (9). The initial damage, D_0 , to be used as the initial condition in Eqs (10) or (11) is, in general, a random variable that takes into account the effects of residual stresses, surface roughness, loading histories etc.

It should be noted that the above formulation of damage growth admits negative damage increments. The probability of such negative damage increment over a given time (or strain) interval depends on the length of the interval, and on the relative magnitude of the drift and diffusion terms. Local and transient retardation in damage might actually occur at the microscale. Nevertheless, the increment of damage should for all practical purposes be non-negative over a finite interval of time and space, in the absence of repair or autogenous healing. This property should be verified in every situation where the model is applied.

For uniaxial monotonic loading, the free energy per unit volume is,

$$\psi = \int \sigma d\varepsilon - \gamma \quad (12)$$

where γ denotes the energy of formation of discontinuities per unit volume due to damage growth. Assuming that (i) the discontinuities are microscopic spheres of different sizes which do not interact with each other, (ii) the force-displacement relation is linear at the microscale, and (iii) stress amplification effects can be neglected, γ can be estimated as (Bhattacharya and Ellingwood, 1998c):

$$\gamma = \frac{3}{4} \sigma_f D \quad (13)$$

where σ_f is the true failure stress. The first term in Eq (12) can be evaluated from the constitutive relation (between the effective stress and strain or strain rate) relevant for the given loading situation.

4. TIME-DEPENDENT RELIABILITY

In this section, the CDM-based stochastic damage growth law derived above is applied to random ductile, fatigue and creep damages, and time-dependent reliability is analyzed in each case.

4.1 Ductile Deformation

The relation between effective stress and total strain under uniaxial monotonic loading may be defined by the Ramberg-Osgood law, $\varepsilon = \bar{\sigma} / E + (\bar{\sigma} / K)^M$, which decomposes the total strain, ε , into its elastic (ε_e) and plastic (ε_p) components, with parameters E = the elastic modulus, K and M = the hardening modulus and exponent, respectively. It is assumed that the exponent M is unaffected by damage. The damaged moduli are $\tilde{E} = E(1 - D)$, $\tilde{K} = K(1 - D)$ for $\varepsilon_p \geq \varepsilon_0$ where ε_0 is the threshold plastic strain for damage initiation (Lemaitre, 1985). The SDE of random ductile damage growth then becomes:

$$dD(\varepsilon_p) = \frac{\varepsilon_p^{1/M} (1 - D(\varepsilon_p))}{\varepsilon_p^{1+1/M} / (1 + 1/M) + C} d\varepsilon_p + \frac{(\sqrt{c_2} / c_1) / K}{\varepsilon_p^{1+1/M} / (1 + 1/M) + C} dW(\varepsilon_p) \quad (14)$$

where $C = [(3/4)(\sigma_f / K) - \varepsilon_0^{1+1/M}] / (1 + 1/M)$. Eq (14) contains two simplifications: (i) $d\varepsilon / d\varepsilon_p \approx 1$, which is true for all ε of interest in ductile deformation damage, and (ii) $K / (2E) \approx 0$, which is valid for most engineering alloys. Eq (14) is of the form of a time-dependent Ornstein-Uhlenbeck process, and since the diffusion term is independent of D , its Ito and Stratonovich solutions are identical (Gardiner, 1985):

$$D(\varepsilon_p) = 1 - (1 - D_0) \frac{(3/4)(\sigma_f / K)}{\varepsilon_p^{1+1/M} / (1 + 1/M) + C} + \frac{(\sqrt{c_2} / c_1) / K}{\varepsilon_p^{1+1/M} / (1 + 1/M) + C} [W(\varepsilon_p) - W(\varepsilon_0)] \quad (15)$$

where $D_0 = D(\varepsilon_0)$ is the initial damage.

| Parameter | Nominal | Mean | C.O.V. | Distribution |
|------------------|----------|---------|--------|---------------|
| E | 74.5 Gpa | - | - | - |
| K | 680 MPa | 680 MPa | 0.20 | Normal |
| M | 5.5 | 5.5 | 0.20 | Normal |
| σ_f | 435 MPa | 435 MPa | 0.20 | Normal |
| ε_0 | 0.016 | 0.016 | 1.0 | Lognormal |
| D_0 | 0 | - | - | Deterministic |
| $\sqrt{c_2/c_1}$ | 20 MPa | - | - | Deterministic |
| D_c | 0.23 | 0.23 | 0.10 | Normal |

Table 1: Material properties for 2024-T3 Aluminum

The random ductile deformation model is validated in Figure 1 with experimental results from Woo and Li (1993) and Lemaitre (1985) for ductile damage growth in 2024-T3 Aluminum. Damage growth data from Lemaitre (1985) however, does not contain any statistical description. The nominal values and statistical properties of the variables are listed in Table 1 (for sources, see Bhattacharya and Ellingwood, 1998a). The material properties, $\ln(\varepsilon_0)$, σ_f , K and M are considered random with moderate stochastic dependence among them (the off-diagonal terms of the correlation matrix are all taken to be 0.5). The means of the random variables are assumed equal to their nominal values. The initial damage is assumed zero since the experiments were carried out on undamaged specimens. The noise intensity $\sqrt{c_2/c_1}=20\text{MPa}$, which is related to the ratio of the variance and the correlation length of the fluctuating quantity s_b , was selected to model the overall magnitude of the observed standard deviation of damage (Woo and Li, 1993). The predicted mean and standard deviation functions of damage in Figure 1 are obtained numerically from Eq (15). The sensitivity to correlation among the random variables and to the noise intensity have been investigated in Bhattacharya and Ellingwood (1998b).

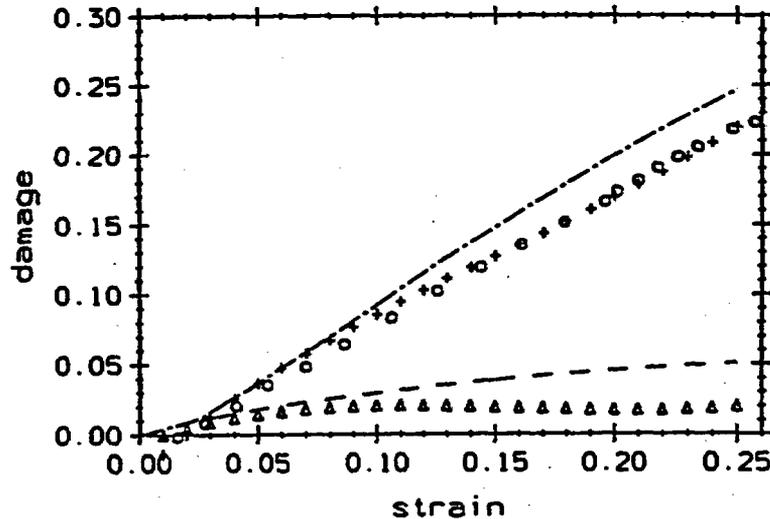


Figure 1 : Random ductile deformation damage

Failure occurs when damage reaches the critical value D_c . If the damage growth rate is almost always positive, the sample paths of $D(\varepsilon)$ which cross D_c from below for the first time may be expected to exist

above that barrier after a finite interval of strain. In such cases, the cumulative failure probability (CFP) can be simplified as the complement of the CDF of the damage function evaluated at the critical damage:

$$F_{\varepsilon_f}(\varepsilon) = 1 - P[D(\varepsilon') \leq D_c; \forall \varepsilon' \in [0, \varepsilon]] \approx 1 - P[D(\varepsilon) \leq D_c] \quad (16)$$

where ε_f denotes the random failure strain.

Figure 2 illustrates the limit state probability for 2024-T3 Aluminum, in which D_c is treated as a random variable (Table 1). Parameters $\ln(\varepsilon_0)$, σ_f , K and M are considered random as before (Table 1), with correlation coefficient 0.5 between each pair. The noise intensity, $\sqrt{c_2/c_1} = 20\text{MPa}$. The sample functions of $D(\varepsilon)$ are obtained numerically from Eq (14) using an interval size $\Delta\varepsilon = 0.01$. None of these randomly selected sample function returns to the safe region once it has exited that region, reinforcing the notion of non-negative damage growth. The relation between D and ε in Eq (16) is not explicit, and the CDF of ε_f is obtained numerically. The mean and standard deviation of ε_f are computed to be, respectively, 0.247 and 0.052, comparable to generally observed values for engineering metals (e.g., Davis, 1993).

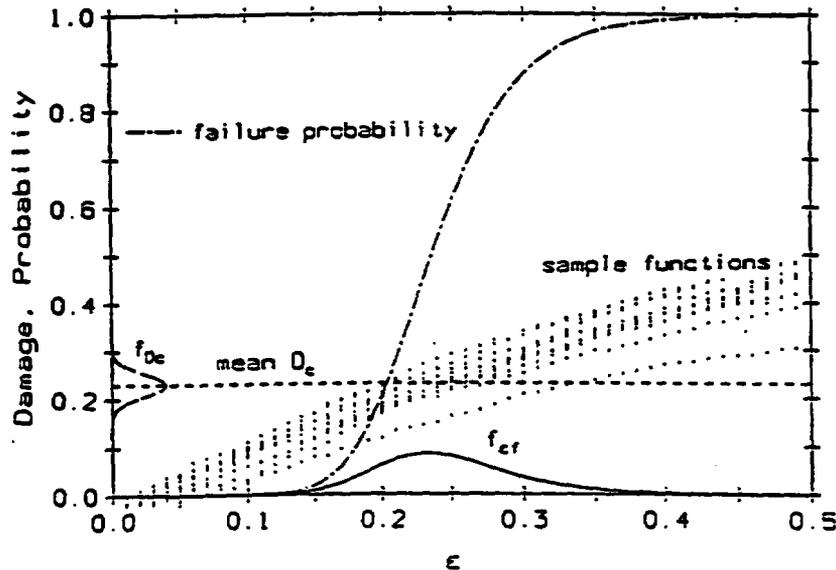


Figure 2: Failure probability and sample paths of random ductile damage

4.2 Fatigue

The total fatigue life, N_T , of a structural member generally consists of two phases: a crack initiation phase of duration N_I , followed by crack propagation phase of duration N_P , such that, $N_T = N_I + N_P$. Depending on past and future loading conditions, the initiation life can be a significant portion of the total fatigue life of a virgin material.

Fatigue damage growth occurs incrementally as a result of load cycling. The damage at the end of cycle i acts as the initial damage for the increment in cycle $i+1$:

$$D_{i+1} = D_i + \Delta D_i, \quad \Delta D_i \geq 0, \quad i = 1, \dots, N_i - 1 \quad (17)$$

We assume that the unloading portion of a hysteresis loop and compressive stresses do not contribute to damage growth so that damage grows only during loading above the endurance limit, S_e , in the positive stress region. Crack initiation occurs when damage exceeds the critical damage:

$$\begin{aligned} D_{N_i-1} &< D_C \\ D_{N_i} &\geq D_C \end{aligned} \quad (18)$$

Applying Eq (10) to fatigue damage growth in cycle i ,

$$\frac{dD}{d\varepsilon} = \begin{cases} -\frac{\sigma_\infty}{\psi_D} + \frac{\sqrt{c_2/c_1}}{\psi_D} \xi(\varepsilon), & \sigma_\infty \geq S_e, \dot{\varepsilon} \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

with the initial condition $D = D_{i-1}$.

The constitutive model for fatigue damage is defined by the cyclic Ramberg-Osgood law with parameters E, K' and M' , which must be obtained from a stabilized cyclic stress-strain curve (Dowling, 1993). The SDE can be solved for fatigue damage growth similarly as in ductile damage, and damage at the end of cycle i is:

$$D_i = 1 - a_i(\varepsilon; \Omega)(1 - D_{i-1}) + b_i(\varepsilon; \Omega)\Delta W_i \quad (20)$$

where a_i and b_i are cycle dependent functions involving strain limits, and ΔW_i is the Wiener increment in cycle i . The recursive nature of the above equation makes it possible to express damage, D_n , at the end of n cycles in terms of the initial damage, D_0 , and n independent increments of the standard Wiener process:

$$D_n = 1 - (1 - D_0) \prod_{i=1}^n g_i + C_0 \sum_{i=1}^n \Delta W_i \prod_{j=1}^i g_j \quad (21)$$

In some situations it may be more convenient to express damage as a function of time, rather than number of cycles. In this case the functional form of $n(t)$ (including its stochastic characteristics) must be incorporated in Eq (21). Assuming that the damage growth process described by Eq (19) is almost always positive, the cumulative probability of failure is the complement of the CDF of D_n ,

$$P[N_i \leq n] \approx P[D_n > D_C] \quad (22)$$

from which the probability distribution of N_i may be obtained if the statistics of D_n and D_C are known.

No published data on random fatigue damage growth (during the pre-initiation stage) oriented toward CDM analysis could be located. Limited data are available on random crack initiation life which allow partial validation of the present model. Predictions of crack initiation in Type A 106-B carbon steel subjected to fully-reversed strain controlled cycling at 288°C in air are compared with experimental

results in Figure 3. The nominal material properties (from Chopra et al 1995) are $E=196.5\text{GPa}$, $K'=1994\text{MPa}$, $M'=7.74$, $\sigma_f=539\text{MPa}$, $\sigma_s=301\text{MPa}$, and $S_c=310\text{MPa}$. The value of the noise parameter, $\sqrt{c_2/c_1}=1000\text{MPa}$, is derived from Keisler et al (1994) to match the standard deviation observed in fatigue tests. The initial damage is treated as zero (deterministic), and the nominal value of D_c is taken as 0.25, which is comparable to the values reported for other carbon steels in Lemaitre (1992). Parameters E , K' , M' , σ_f , S_c and D_c are considered random and statistically independent of each other. The mean values of these six random variables are assumed equal to their respective nominal values. The first five random variables are assumed lognormal and D_c is assumed normal. All six are assumed to have a c.o.v. (coefficient of variation) of 10%. The nominal strain ratio is $R = -1$ (fully reversed cycling), and the ordinate of Figure 3 represents the nominal values of the strain amplitudes, $\Delta\epsilon/2$. The strain amplitudes are considered statistically independent and identically distributed lognormal random variables, each having c.o.v. 10%. The predicted mean initiation time, $\mu(N_i)$, along with the bound of one standard deviation above mean, $\mu(N_i) + \sigma(N_i)$, compares well with the (i) estimated N_i from Majumdar et al (1993) which corresponds to the formation of a 0.18mm crack, (ii) the cycles to failure, N_F from Majumdar et al (1993), and (iii) N_{25} , from Chopra et al (1995) and Chopra (1996) which correspond to a 25% drop in the peak stress and a 3mm crack.

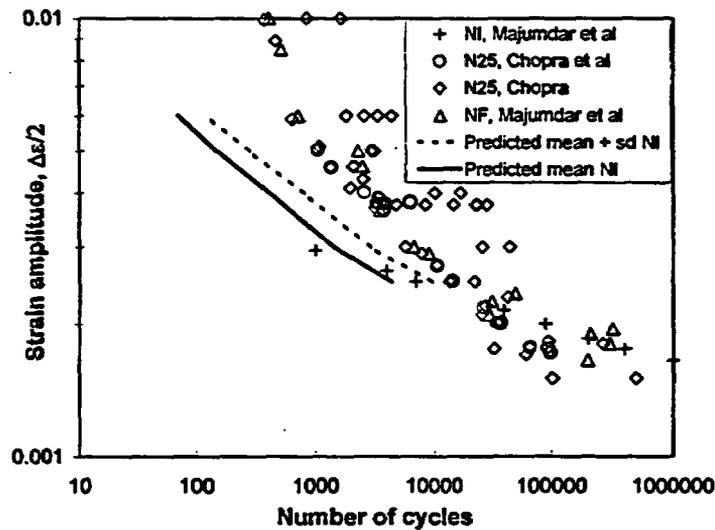


Figure 3 : Random fatigue damage growth in A106 Gr B steel at 288C in air

4.3 Creep

The equivalence principle [cf Eqs (1) and (2)] applied to creep strain rate, $\dot{\epsilon}_{cr}$, as given by the Bailey-Norton law under uniaxial loading (Dowling, 1993), gives:

$$\dot{\epsilon}_{cr} = A\phi\tilde{\sigma}^m t^{\psi-1} \quad (23)$$

where $\tilde{\sigma}$ is the effective applied stress. Under constant stress creep, $\dot{\epsilon} = \dot{\epsilon}_{cr}$, and ψ_D can be simplified as $\psi_D = -(3/4)\sigma_f$, where σ_f is the true failure stress at the operating temperature. The stochastic differential equation of isotropic creep damage growth thus becomes,

$$dD(t) = \frac{A_1}{(1-D(t))^m} dt + \frac{B_1}{(1-D(t))^m} dW(t) \quad (24)$$

where,

$$A_1 = \frac{4}{3} \frac{A \phi t^{\phi-1}}{\sigma_f} \sigma_\infty^{m+1} \quad (25)$$

$$B_1 = \frac{4}{3} \frac{A \phi t^{\phi-1}}{\sigma_f} \sigma_\infty^m \frac{\sqrt{c_4}}{c_3}$$

To the knowledge of the authors, Eq (24) does not have a closed-form solution in the Ito sense. But a closed-form solution is possible in the Stratonovich sense, under the condition $\phi=1$ (steady state creep):

$$D(t) = 1 - \left\{ (1 - D_0)^{m+1} - A_1 (m+1)t \right\}^{1/(1+m)} \cdot \left[1 - \frac{B_1 (m+1)W(t)}{(1 - D_0)^{m+1} - A_1 (m+1)t} \right]^{1/(1+m)} \quad (26)$$

where the initial time $t_0=0$. The initial damage, D_0 , accounts for the ductile damage caused when the component is loaded to σ_∞ at the beginning of the process, in addition to any damage existing prior to the commencement of creep straining. Eq (26) has the same form as the deterministic solution for steady state creep (Bhattacharya and Ellingwood, 1999), with an additional term (in square brackets) containing noise.

Creep damage has an accelerated growth rate with respect to time (e.g., Kachanov, 1986) in the pre-localization stage. Damage, however, is bounded in the range $[0, D_c]$ by definition, and $D_c \leq 1$ from physical considerations. Sample functions of $D(t)$ are absorbed by the boundary $D = 1$. The cumulative distribution function (CDF) of $D(t)$, is therefore a mixed distribution in $[0, 1]$. Assuming that creep damage growth rate is almost always positive, the cumulative failure probability in the interval $[0, t]$ is,

$$F_{T_f}(t) = 1 - P[D(\tau) \leq D_c; \forall \tau \in [0, t]] \approx 1 - P[D(t) \leq D_c] \quad (27)$$

where T_f is the random time to failure.

No published CDM-based studies (e.g., by measuring the reduced stiffness) of stochastic creep damage growth could be located. Therefore, in validating the proposed random creep damage growth model, only the statistics of the predicted failure time are compared with available experimental results. The material chosen is type 316 stainless steel stressed to 199MPa at 593°C (1100°F). The nominal creep law and tensile parameters for type 316 stainless steel at 593°C are listed in Table 2 (Davis, 1994; Garofalo et al, 1961). Parameters D_0 , A , m , D_c are considered as random and mutually statistically independent. As before, the mean value of a parameter is taken equal to its nominal value. The nominal value of D_0 is computed from Eq (15) with deterministic E , K , σ_f , M , zero prior damage and zero noise.

Figure 4 shows the predicted mean and standard deviation of damage and the failure probability as functions of time, under the Stratonovich interpretation [Eq (26)]. The mean and c.o.v. of the failure time are 924 hr and 43% respectively. These values may be compared with the scatter observed by Garofalo et al (1961) in the times (hours) to (i) the onset of tertiary creep (mean=1283, c.o.v.=0.28, min=960, max=1950) and (ii) rupture (mean=1749, c.o.v.=0.21, min=1267, max=2437); under the same conditions of temperature and stress. Comparison between Ito and Stratonovich solutions were performed in Bhattacharya and Ellingwood (1998a).

| Parameter | A MPa, hr | m | ϕ | test σ MPa | σ_f MPa | E GPa | K MPa | M | D_0 | D_c | $\sqrt{c_1/c_3}$ MPa $\sqrt{\text{hr}}$ |
|---------------|------------------------|------|--------|----------------------|-------------------|------------|------------|------|--------|-------|---|
| Nominal | 2.32×10^{-20} | 6.92 | 1 | 199-315 | 443.7 | 151.6 | 492.7 | 4.22 | 0.0108 | 0.20* | 300* |
| C.O.V.* | 0.20 | 0.01 | - | - | - | - | - | - | 0.10 | 0.10 | - |
| Distribution* | LN | N | det | det | det | det | det | det | N | N | det |

Table 2: Creep and tensile properties of type 316 stainless steel at 593°C
(*=assumed, LN=lognormal, N=normal, det=deterministic)

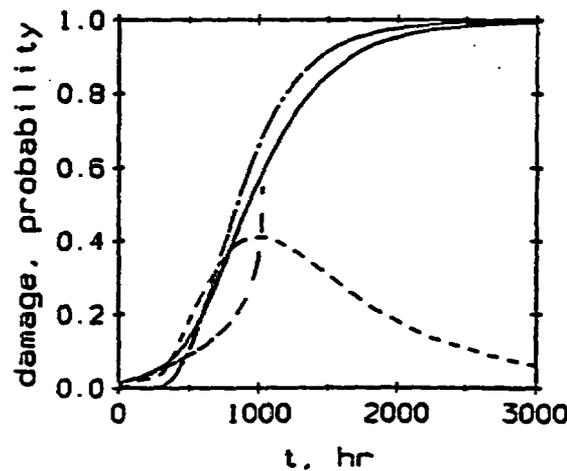


Figure 4: Creep damage growth statistics in type 316 stainless steel
[——— $\mu_D(t)$, - - - - $\sigma_D(t)$, - · - · $F_T(t)$, · · · · deterministic $D(t)$]

The creep damage accumulation model and time-dependent reliability analysis are illustrated with an application to an aging and corroding cylindrical steel pressure boundary subject to a sequence of severe operating events involving pressure and temperature. This pressure boundary is assumed to be designed by ASME requirements. Table 3 summarizes the parameters used in the illustration.

The number of pressure/temperature occurrences, $N(t)$, is modeled as the sum of two independent Poisson point processes, $N(t) = N_0(t) + N_A(t)$, where (i) $N_0(t)$ is a pure "chance" phenomenon like human error, with constant mean rate λ_0 , and (ii) $N_A(t)$, with mean rate $\lambda_A(t)$, represents an "aging" phenomena that might cause safety systems to malfunction. $\lambda_A(t)$ is given by $\lambda_A(t) = (\alpha/u)(t/u)^{\alpha-1}$, in which u, α are parameters, consistent with the common assumption that failure times are described by a Weibull distribution. Values of α greater than 1 represent a realistic "aging" process.

| Variable | Nominal (design) value | Statistical properties (mean, c.o.v.) |
|----------------------------------|------------------------|---------------------------------------|
| Original thickness, h_0 | 1.375 in (34.9 mm) | deterministic |
| Peak pressure, P_m | 60 psig (0.42MPa) | Type I max (0.8 $P_{des,n}$, 20%) |
| Peak temperature, θ_m | 390°F (199°C) | Type I max (177°C, 30%) |
| Significant duration, Δt | 20 min | Lognormal (1000s, 60%) |
| Yield stress, F_y | 38 ksi (262 Mpa) | Lognormal (1.10 F_m , 7%) |

Table 3 : Original (uncorroded) dimensions, load and strength statistics

During a serious operating event, it is assumed that the temperature and pressure rise in a very short time to their peak values, P_m and θ_m respectively, and remain constant at the peak values for a duration of Δt . Creep damage is idealized to occur over Δt at constant temperature θ_m under the action of the constant load P_m . The action of the load appears in the form of the membrane stress, which is aggravated by corrosion loss. Statistical dependence may exist between P_m and θ_m during the operating event.

The random penetration, $Z(t)$, of uniform corrosion is modeled by,

$$Z(t) = C(t - T_i)^M, t \geq T_i \quad (28)$$

in which C = random rate parameter, M = random time-order parameter, and T_i = random initiation period. The corrosion degradation process is assumed to occur slowly enough that the time-dependent resistance variables can be treated as constants during the duration of pressurization (Ellingwood and Mori, 1993).

Suppose that n events of duration Δt_i occur at random instants of time, t_i ($i = 1, 2, \dots, n$). The accumulated creep damage, D_n , is given by (Bhattacharya and Ellingwood, 1998c):

$$(1 - D_n)^{m+1} = (1 - D_0)^{m+1} - \sum_{i=1}^n A_i \sigma_i^{m+1} \Delta t_i - \sum_{i=1}^n B_i \sigma_i^m W(\Delta t_i) \quad (29)$$

where the coefficients A_i and B_i depend on the temperature, θ_{mi} ; membrane stress, σ_i , depends on the pressure, P_{mi} , and the remaining shell thickness, $h_0 - Z(t_i)$. Figure 5 shows several sample functions of creep damage, providing a schematic representation of creep damage accumulation during a series of severe operating events. The creep parameters are (in ksi, hr): $A=1.0 \times 10^{-10}$, $m=5$, $\phi=1$ and $\sqrt{c_4/c_3}=10$. The corrosion parameters are: $M \sim \text{Normal}(0.7, 20\%)$, $T_i \sim \text{Lognormal}(10 \text{ yr}, 30\%)$ and $C \sim \text{Lognormal}(230 \mu\text{m}, 30\%)$. The load process is given by $\lambda_0 = 0.1/\text{yr}$, $\alpha = 3$ and $u = 25 \text{ yr}$, and is assumed to approximate a severe operating event history. A correlation coefficient of 0.6 has been assumed between P_m and θ_m . Removing the conditioning on n results in an unconditional estimate of creep damage accumulation as a function of time.

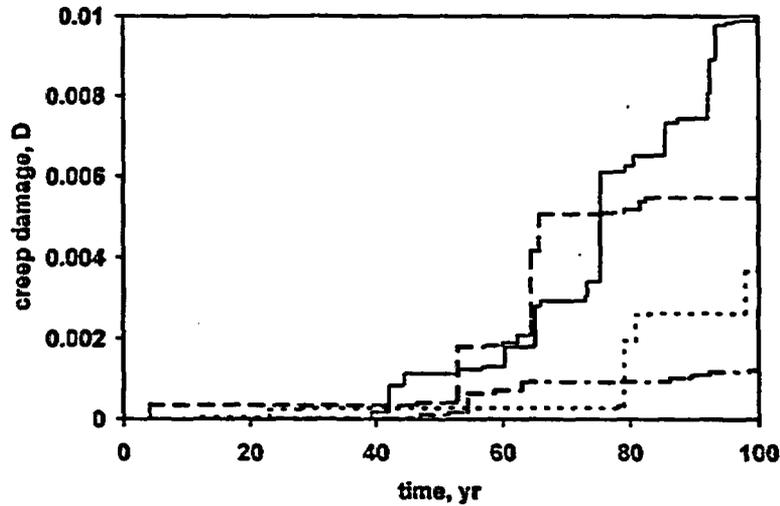


Figure 5 : Schematic of creep damage accumulation in presence of aging and corrosion

4.4 Corrosion

It has been shown in Bhattacharya and Ellingwood (1998c) that in the presence of corrosion, the failure probability due to creep damage accumulation is at least an order of magnitude lower than the probability of failure due to excessive inelastic deformation. The treatment of corrosion above was accomplished by an empirical rate equation not connected with CDM. It is theoretically possible to cast corrosion damage caused by oxidation, carbonation etc in a CDM format, as mentioned by Cauvin and Testa (1999) in the context of fourth order damage tensors. The assumption of isotropic damage no longer holds when corrosion is the major cause of damage, since corrosion is a surface phenomenon. The effective stress is enhanced in the presence of corrosion, and the free energy needs to be suitably modified to account for the relevant chemical reaction.

5. CONCLUDING REMARKS

Time-dependent reliability analyses have been performed for steel elements and steel pressure boundary components subjected to corrosion, ductile damage from sustained load, elevated temperature creep, and low-cycle fatigue (Bhattacharya and Ellingwood, 1998c). The CDM approach was validated for limit states involving ductile damage, creep and low-cycle fatigue. It was found that corrosion has the most significant impact on time-dependent reliability of steel components, the other mechanisms having a more localized effect. The estimated conditional failure rates for structural components increase in a nonlinear fashion with time. Neglecting this nonlinear behavior may lead to an erroneous appraisal of time-dependent margins of safety.

Acknowledgments

Support for the research, provided, in part, by Lockheed-Martin Energy Research Corporation under Grant 19X-SP638V, with Dr D. J. Naus as Program Manager, is gratefully acknowledged. The authors would also like to thank Mr. Omesh Chopra of Argonne National Laboratory for making available the fatigue loading data used in this paper.

References

- Bhattacharya, B. and Ellingwood, B. R. (1999). "A New CDM-based approach to structural deterioration", *International Journal of Solids and Structures*, 36(12):1757-1779.
- Bhattacharya, B. and Ellingwood, B. (1998a). "Continuum damage mechanics-based model of stochastic damage growth", *Journal of Engineering Mechanics, ASCE*, 124(9):1000-1009.
- Bhattacharya, B. and Ellingwood, B. R. (1998b). "A CDM analysis of stochastic ductile damage growth and reliability", *Journal of Probabilistic Engineering Mechanics*, 14(1-2):45-54
- Bhattacharya, B. and B. Ellingwood (1998c). "A damage mechanics based approach to structural deterioration and reliability." NUREG/CR-6546, US Nuclear Regulatory Commission, Washington, DC.
- Callen, H. B. (1985). *Thermodynamics and an Introduction to Thermostatistics*, Wiley.
- Cauvin, A. and Testa, R. B. (1999). "Damage mechanics: basic variables in continuum theories", *International Journal of Solids and Structures*, 36:747-761.
- Chaboche, J. L. (1988). "Continuum damage mechanics - I and II", *Journal of Applied Mechanics*, 55:59-72.
- Chopra, O. K. et al (1995). Environmentally Assisted Cracking in Light Water Reactors}. Report NUREG/CR-4667. U.S. Nuclear Regulatory Commission, Washington, DC.
- Chopra, O. K. (1996). Private communication to B. R. Ellingwood (author of this paper).
- Chow, C. L. and Wei, Y. (1991). "A model of continuum damage mechanics for fatigue failure", *International Journal of Fracture*, 50(4):301-316.
- Davis, J. R. (1993). *Aluminum and Aluminum Alloys*, ASM International, Materials Park, OH.
- Davis, J. R. (1994). *Stainless Steels*, ASM International. Materials Park, OH.
- Dowling, N. E. (1993). *Mechanical Behavior of Materials*, Prentice Hall.
- Ellingwood, B., Bhattacharya, B. and Zheng, R.-H. (1996). "Reliability-based condition assessment of steel containments and liners." NUREG/CR-5442, US Nuclear Regulatory Commission, Washington, DC.
- Ellingwood, B.R. and Mori, Y. (1993). "Probabilistic methods for condition assessment and life prediction of concrete structures in nuclear power plants", *Nuclear Engineering and Design*, 142:155-166.
- Gardiner, C. W. (1985). *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer-Verlag.

- Garofalo, R. W. et al (1961). "Creep and creep rupture relationships in an austenitic stainless steel", *Transactions of the Metallurgical Society of AIME*, 221: 310-319.
- Kachanov, L. M. (1986). *Introduction to Continuum Damage Mechanics*, Martinus Nijhoff.
- Keisler, J., Chopra, O. K. and Shack, W. J. (1994). *Statistical Analysis of Fatigue Strain-Life Data for Carbon and Low-Alloy Steels*, report NUREG/CR-6237, ANL-94/21. Argonne National Lab., Argonne, IL.
- Krajcinovic, D. (1984). "Continuum damage mechanics", *Applied Mechanics Reviews*, 37(1):1-6.
- Krajcinovic, D. and Sumarac, D. (1987). "Micromechanics of the damage process", *Continuum Damage Mechanics Theory and Applications*, eds: D. Krajcinovic and J. Lemaitre, Springer-Verlag.
- Lemaitre, J. (1992). *A Course on Damage Mechanics*, Springer-Verlag.
- Lemaitre, J. (1985). "A continuum damage mechanics model for ductile fracture", *Journal of Engineering Materials and Technology*, 107(1):83-89.
- Majumdar, S., Chopra, O. K. and Shack, W. J. (1993). Interim failure design curves for carbon, low-alloy, and austenitic stainless steels in LWR environments. *Proceedings 20th WRSM*, vol 3, NUREG/CP-0126, vol 3. March 1993.
- Naus, D. J., Oland, C. B. and Ellingwood, B. (1996). "Report on aging of nuclear power plant reinforced concrete structures." NUREG/CR-6424, US Nuclear Regulatory Commission, Washington, DC.
- Oland, C. B. and Naus, D. J. (1998). "A survey of repair practices for nuclear power plant containment metallic pressure boundaries." NUREG/CR-6615, US Nuclear Regulatory Commission, Washington, DC.
- Ostojca-Starzewski, M. (1989). "Damage in a random microstructure: size effects, fractals and entropy maximization", *Applied Mechanics Reviews*. 42(11):S202-S212.
- Woo, C. W. and Li, D. L. (1993). "Statistical analysis of material damage with changing internal structure." *Engineering Fracture Mechanics*, 45(2):245-254.

Feasibility Study on the Use of Ultrasonic Technology on Embedded Corrosion Detection of Nuclear Containment Units, Phase II

Final Report

July 17, 1998

Prepared by
Jason Rudzinsky
Matt Conti
Joe Bondaryk

Cambridge Acoustical Associates/Engineering Technology Center
84 Sherman Street
Cambridge, MA 02140

Abstract

The nuclear power industry is concerned with corrosive thinning of containment unit sections embedded in concrete. This study investigated the feasibility of detecting these thickness degradations using ultrasonic imaging. A commercial ultrasonic system was used to carry out several full-scale, controlled, laboratory experiments. Measurements of 0.5MHz shear wave levels propagated in one inch thick steel plate embedded in concrete showed 1.6dB of signal loss for each centimeter of two way travel in the steel plate (compared to previous numerical predictions of 3-4dB). Negligible losses were measured in plates with a decoupling treatment applied between the steel and concrete. Scattered signals from straight slots of different size and shape were investigated. The return from a 4mm deep rectangular slot exhibited levels 24dB down relative to incidence and 4-6 dB higher than those obtained from both "v" shaped and rounded slots of similar depth. The system displayed an input/output dynamic range of 125dB and measurement variability less than 1-2dB. Based on these results, a 4mm deep, rounded degradation embedded in 30cm of concrete has expected returns of -76 to -78dB relative to the input and should therefore be detectable. (Work is supported by the Oak Ridge National Laboratory and the Nuclear Regulatory Commission.)

1.0 Introduction

The Nuclear Regulatory Commission has a program in conjunction with Oak Ridge National Laboratories, to investigate structural monitoring of aging nuclear containment units. Engineering Technology Center was commissioned as a subcontractor to investigate the feasibility of employing ultrasonic imaging technologies to the problem of detecting corrosive degradations in embedded or inaccessible regions of containment units. The work was sponsored in two phases. The first phase addressed basic feasibility issues using numerical models. The second phase, which is the focus of this report, utilized experimental means to continue the feasibility study as well as verify the phase

one findings.

1.1 Statement of Problem and Solution Approach

An area of concern in the nuclear power industry is the structural integrity of aging and inaccessible containment unit sections that are embedded in concrete. One of the fears is that over time, water intrusion causes corrosive thinning and pitting in inaccessible areas of the pressure vessel and may go undetected (see Figure 1). Other than expensive and potentially dangerous concrete chipping techniques, there is no procedure currently in place for detecting degradations in these regions.

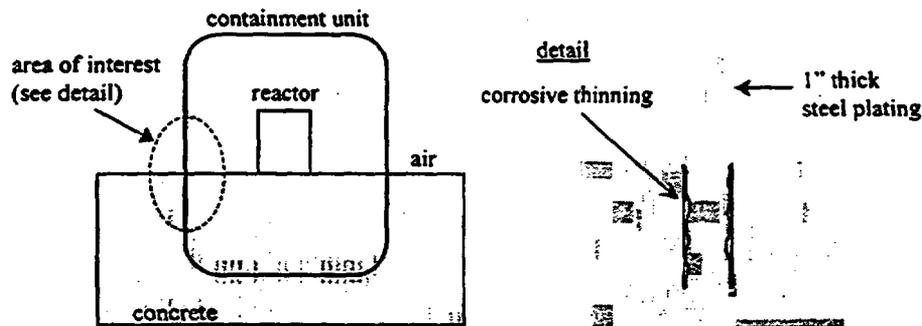


Figure 1 - Overview of problem.

The overall objective of this research project has been to study the feasibility of employing alternate ultrasonic imaging technologies for the detection and localization of degradations in embedded regions. The focus has been on studying high frequency (0.5 MHz - 2.5 MHz) structural waves propagated laterally from accessible regions of a free-standing steel pressure vessel (as opposed to a steel lined, concrete vessel) to degradations below the air-concrete interface. The reflected returns from these structural waves can be processed to generate an image of the degradation.

The basic approach differs from a conventional ultrasonic thickness (UT) or "through thickness" test. During a through thickness test, a transducer is placed in direct contact with a test structure (see Figure 2) and pulse-like waves are injected normal to the test structure's surface. The time delay between sending and receiving a pulse, combined with an assumed knowledge of the test material's compressional sound speed provide means for estimating the material's local thickness. This technique is used in many industrial applications and has attained a relatively high level of refinement. However, at a minimum, the area to be tested must be accessible, which is clearly not the case for the problem at hand.

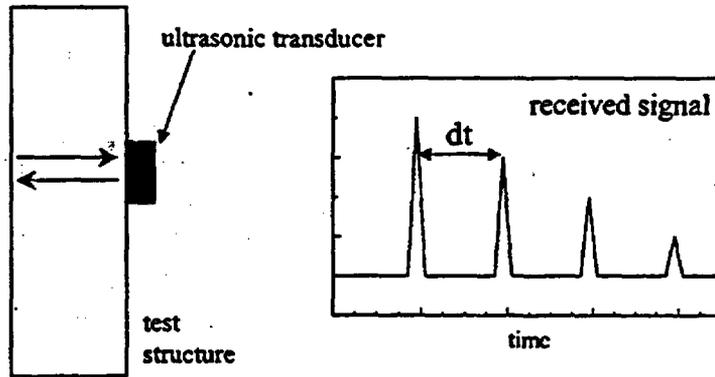


Figure 2 - Procedure used in conventional UT test.

The proposed alternate technique, known as an Angle Beam Inspection, instead uses a plastic wedge to couple the transducer to a test structure (see Figure 3). The transducer generates compressional waves in the wedge that are refracted, primarily as shear waves, in the test structure. These refracted waves then "skip" laterally away from the source through the structure. The technique is often used to inspect welded joints and to determine the presence of cracks or other structural flaws. The incident waves for this type of test are not typically required to propagate over considerable distances or to propagate in constrained regions of the test structure, which they will be required to do in order to demonstrate success in the proposed test scenarios.

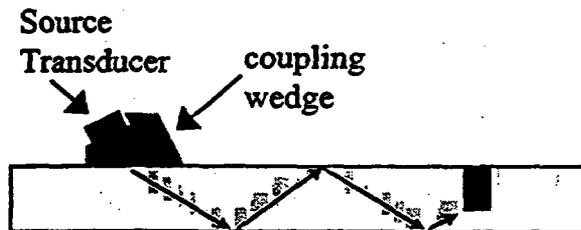


Figure 3 - Procedure used in Angle Beam Inspections.

The feasibility of employing the technique under the proposed scenario therefore centers on determining if a measurable and decipherable signal is returned from the corrosion area. This study addresses the physics side of the problem by investigating the energy lost from the interrogating signals into the surrounding concrete, as well as characterizing corrosive type degradations as acoustic scatterers. Also of primary importance is the practical issue of determining the performance limits exhibited by a commercially available Angle Beam Inspection System. The ability to discern and extract reflector characteristics using acoustic signals that include information about both the corrosive damage and the propagation path is an important, but at this point, secondary concern.

1.2 Review of Previous Work

The present work builds on a previous theoretical feasibility study which utilized a widely accepted elastic layered media numerical computer code (OASES) to model the angle beam inspection scenario. Detailed analyses and conclusions can be found in Reference 1. The major conclusions derived from the numerical study include the following:

- The embedding concrete introduces 3-4 dB of signal loss for each centimeter of two-way travel in embedded plates. The “non-spreading” propagation losses in the free-standing steel portions are negligible.
- Notch degradations 4mm in depth across the plating thickness display reflection coefficients (ratio of reflected to incident wave amplitude including removal of geometric spreading losses) of roughly -23dB at 0.5 MHz.
- Variation in backscatter is weakly dependent on degradation depth (so long as the interrogating wavelength is greater than twice the degradation depth), frequency and wedge angle.
- The technique does not appear to be applicable to steel lined, concrete vessels due to an immense loss of energy into the concrete shell.

The study encouragingly showed that with sufficient, albeit demanding, input/output measurement system dynamic range, the anticipated return levels from a representative degradation located in an embedded region of a steel pressure vessel should be detectable.

1.3 Objectives for Present Work

In an effort to not only provide a basis for improving the numerical models, but also to continue the feasibility study in a more practical forum, a series of controlled laboratory experiments were designed and comprise the majority of the efforts contained in this report. This work represents the second phase of the project.

The project's objectives can be broken down into two main areas. The first area concerns characterizing the overall performance limits of a commercially available ultrasonic measurement system. These characterizations will include:

- Determining the system's input/output dynamic range
- Determining measurement stability and repeatability
- Determining the source quality
- Determining the noise environment.

The system's dynamic range will dictate the amount of loss that can be incurred in a given scenario. In an effort to estimate the total loss that will be induced on an incident signal, several building blocks that can be used as components of total loss will be quantified experimentally. The following mechanisms were assumed to play the most prominent roles in contributing to the total signal loss:

- Transmission past the degradation
 - Degradation shape effects
 - Degradation depth effects
- Geometric spreading
- Additional propagation losses (e.g. waveguide surface interactions)

- Losses to embedding concrete.

Quantifying these mechanisms comprised the second major series of tests. The resultant information provides a means for determining the types of scenarios in which the technique will demonstrate success.

1.4 Organization of Report

This section of the paper has detailed the problem at hand as well as the basic approach to solving the problem. This section also reviewed the previous work's major conclusions and outlined this phase's objectives in determining the feasibility of applying ultrasonic imaging to the problem of embedded corrosion detection. The Phase I results play a role in defining the focus of the Phase II study in that only free-standing steel containment units are studied. Section 2 lays out the background information needed to offer explanations for the experimental results. Section 3 details the experimental results, presenting first a series of experiments meant to study the performance and operability of an Angle Beam Inspection system and then a series of tests designed to quantify the individual signal loss components outlined in Section 1.3. Section 4 summarizes the conclusions obtained from the experimental work and how they affect the technique's feasibility. Finally, Section 5 presents procedural recommendations for implementing the technique and the remaining issues that must be addressed before the technique can be put into practice. The remaining issues primarily focus on experiments that should be carried out to aid in deriving detection, localization, and degradation characterization algorithms.

2.0 Background

2.1.1 Review of Basic Physics - Directivity of Transducer

Before detailing the experimental test procedures, it is useful to first review the basic physics involved with ultrasonic testing. An ultrasonic transducer is comprised of a metallic external housing which contains a piezo-ceramic disk backed by a high-density material. The piezo element deforms as a voltage is applied across opposite faces. As the piezo element is set into motion, a protective face plate attached to one side of the piezo element, is driven and ideally acts as a rigid piston radiator of sound waves into the medium to which it is coupled. The variation of sound wave amplitude as a function of radiating angle is known as the farfield directivity pattern of the source. For a rigid circular piston radiator, the farfield directivity pattern of injected sound waves into a shear free media (the directivity pattern set up in a solid is considerably more complex, but the basic concept holds) is governed by the following equation (Ref. 2):

$$D(\theta) = \frac{J_1[2\pi(a/\lambda)\sin\theta]}{2\pi(a/\lambda)\sin\theta}$$

Here, λ is the acoustic wavelength ($\lambda = c/f$, where c is the speed of propagation and f is frequency) propagation, a is the radiator diameter, and $J_1(x)$ is a Bessel Function of the

first kind, having order 1 and argument x . Figure 4 shows a plot of the farfield sound wave directivity pattern set up by a one-inch diameter radiator vibrating at 0.5 MHz and 1.0MHz (note that the directivity pattern is axisymmetric due to the symmetry of the radiating surface). Several features of the radiators' directivity patterns are to be noted:

- The directivity pattern takes the form of a main "lobe", centered on the surface normal, and several lower level secondary lobes away from the surface normal direction.
- A radiator of fixed diameter displays a more focused field when vibrating at higher frequencies.
- The half power beam width is defined as the angle at which a radiator's directivity factor is -3dB relative to the on-axis (0°) value and is typically used to represent the radiator's spreading angle.

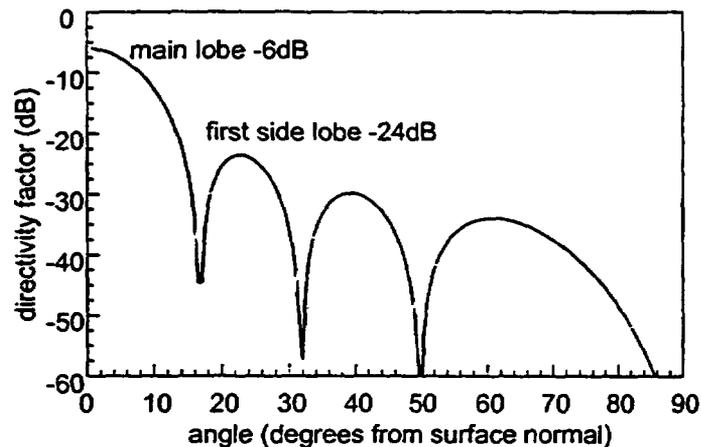


Figure 4 - Directivity factor for a rigid disc radiating in fluid. The black curve represents the relative magnitude radiated as a function of angle (measured from the radiator's surface normal) for a 1" disk vibrating at 0.5MHz and the gray curve a 1" disk vibrating at 1.0MHz.

Close to the transducer, there exists a "nearfield" acoustic response which is considerably more complex than the farfield response and cannot be expressed in such compact form. For relatively low frequency ultrasonic transducers (such as the ones used in the following experiments), the nearfield exists within a one to two inch radius from the transducer. Because the wedges used in the experiments are not sufficiently large to contain the entire nearfield response, the complexity in fully modeling the ultrasonic signals that are injected into a test structure is compounded. For the current study, the farfield approximation of the beam pattern was used. The success of the overall technique is contingent upon a thorough knowledge of the spatial and temporal character of the incident sound waves. Such knowledge is required to properly decipher the returned signals. Consequently, one important recommendation for future work is that the injected source directivity be studied in considerably more detail.

2.1.2 Review of Basic Physics - Transmission through an Interface: Snell's Law

When waves traveling in one medium encounter an interface with a second

medium of different acoustic impedance ($z_{ac} = \rho c$, where ρ is the medium's density, and c is the medium's sound speed, see Figure 5) the wave is partially reflected and partially transmitted. The transmitted wave propagation direction is skewed according to Snell's law:

$$\frac{\sin \theta_1}{c_1} = \frac{\sin \theta_2}{c_2}$$

where θ_1 and θ_2 are noted in Figure 5, and c_1 and c_2 are the wave propagation speed in media 1 and 2, respectively.

In solids, internal waves can be propagated as shear and compressional deformations. Additionally, waves of one type can be converted to another when an impedance discontinuity is encountered. Note that Snell's Law can be rearranged to solve for the refracted transmission angle, θ_2 ,

$$\theta_2 = \sin^{-1}((c_2/c_1)\sin \theta_1)$$

and that for certain material sound speed combinations and certain incident angles θ_1 , there is no obtainable solution. For example, compressional waves cannot be excited in steel ($c_c=5500\text{m/s}$) when compressional waves travelling in Lucite (a material typically used in UT coupling wedges, $c_c=2680\text{m/s}$) are incident at angles greater than 27 degrees. For reference, the shear wave speed in steel is around 3300m/s.

For normal incidence, the ratio of transmitted to incident wave amplitude is governed by the following equation (Ref. 3):

$$\frac{|T|}{|I|} = \frac{2z_2}{z_2 + z_1}$$

As one would expect, when $z_1 = z_2$, all of the wave energy is transmitted. When $z_1 \gg z_2$, the transmitted amplitude is negligible, so virtually all of the wave energy is reflected. Waves travelling from the source wedge (for Lucite, $z_{ac} = 3.16 \times 10^6 \text{ Kg/m}^2\text{s}$) to a steel ($z_{ac} = 45.4 \times 10^6 \text{ Kg/m}^2\text{s}$) test structure are therefore mainly transmitted. Those waves that do get reflected set up a reverberant field inside the wedge, which dies out somewhat quickly, but can still pose problems in detecting small flaws near the source. Waves travelling in steel that encounter an air interface are virtually totally reflected.

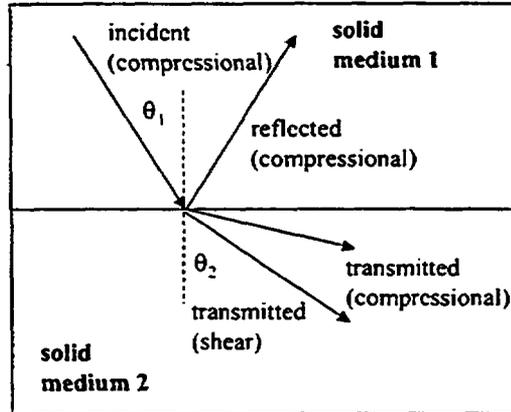


Figure 5 - Wave reflection and transmission at an acoustic impedance interface.

These phenomena are also helpful in describing the structural wave coupling between the steel plate and the embedding concrete. Shear waves in steel have an acoustic impedance five times that of concrete. One can therefore anticipate fairly significant losses from shear waves in steel to compressional waves in embedding concrete. However, these losses are strongly dependent on the interfacial coupling conditions, which are not well known. The transmission equations assume continuity of displacement and stress across the impedance discontinuity, a condition that is not necessarily achieved at the concrete-steel interface. Without an adhesive-like bond, there is minimal interfacial coupling. For this reason, a fluid couplant is used between the transducer-wedge interface and the wedge-plate interface.

2.1.3 Review of Basic Physics - Wave Propagation and Scatter

When a wave is injected into a layered media, e.g. a plate, and the incident wavelength is significantly smaller than the plate thickness, it can be modeled as a propagating ray (see Figure 6). The ray "skips" laterally down the waveguide, bouncing off of the layer's top and bottom surfaces. Locally, where a bounce occurs, the surface displacement is high, giving rise to discrete "hot spots" along the surfaces of the layer. The distance between adjacent hot spots is termed the "skip length" and can be calculated using this simple formula:

$$L_{skip} = 2T \tan \theta$$

where T is the material thickness. Because the wave injected has angular spread, and the wave is free to spread side to side in the layer (or out of Figure 6's page, as seen in Figure 7) there is a geometric spreading loss. The geometric loss in a two dimensional waveguide (such as a semi-infinite plate) is proportional to the inverse of the square root of the distance traveled, and is termed cylindrical spreading.

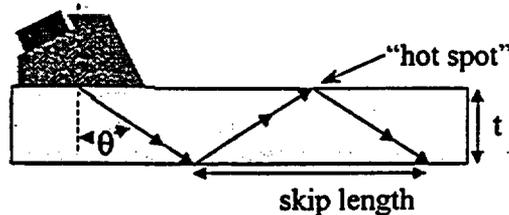


Figure 6 - Basic guided wave propagation in a thick layer

Impedance discontinuities, or degradations of characteristic dimension less than one half of the interrogating wavelength, are inefficient acoustic scatterers. The criteria for detection is that the structural wavelength be small (and the frequency high) compared with the characteristic depth of the degradation. However, the inherent damping loss associated with wave propagation in any media is proportional to frequency, as shorter wavelength signals must endure more cycles to interrogate a fixed distance and therefore suffer a greater loss. Thus for maximum wave penetration in an elastic media, the lowest possible frequency should be selected. Clearly a compromise is required that addresses these two bounding conditions when small degradations are to be detected at considerable distances.

2.2 Discussion of Equipment

Several manufacturers of ultrasonic transducers and peripheral equipment were solicited for information on their products. A completely integrated laboratory testing package (pulsar-receiver electronics, acquisition software and hardware, transducers and wedges) manufactured by Matec Instruments, Inc. (of Northborough, Massachusetts) was selected for use in the experiments.

The testing equipment is comprised of a pulser/receiver card, a high frequency analog-to-digital acquisition board (both of which are attached to a standard PC ISA bus), controlling software, piezo element contact transducers, coupling wedges and industrial grade ultrasonic gel couplant. The transducers have a one inch diameter circular radiating face, are tuned to 0.5MHz, and have an estimated half angle beam width of just over six degrees.

Wedges are specified by the refracted shear wave angle (measured from the surface normal) introduced into a steel test structure. All of the wedge angles that were used (45°, 60°, and 70°) are past the compressional wave critical angle for steel, meaning that compressional waves will not be excited in the steel plate. It should again be emphasized that the transducer does not generate a planar incident field, so that energy is injected in all directions. Therefore, it is more accurate to say the main lobe of the incident sound field does not excite compressional waves in a steel test structure, but that the sidelobes below the main lobe may indeed excite compressional waves.

The equipment can be utilized in either *through transmission* mode, where separate transducers act as source and receiver, or *pulse-echo* mode, in which a single transducer injects a wave and then passively listens for a return (these modes of operation

are known to the acoustics community as *bistatic* and *monostatic* respectively). When operated in through transmission mode, the source and receiver transducers are attached to coupling wedges of the same angle.

The input voltage waveform is a toneburst, with the signal length, level (in percent of maximum) and center frequency specified by the user. The signal length used in all of the experiments was 4 microseconds, and the frequency 0.5MHz. The maximum input voltage that the pulser can generate is 300v. Because the transducers are uncalibrated, it is relationship between input voltage and mechanical force is unknown. It is important to note that several more expensive pulser/receiver cards that can generate inputs of 1000v are commercially available. The acquisition board range is +/- 0.5 volts. Therefore the maximum measurable signal is -6dB referencing 1 volt and the minimum is around -130dB, with the minimum set by the system's dynamic range, which will be shown to be 125dB.

2.3 Discussion of Test Platforms

The test platforms to be used are 1.0"x36.0"x8.0" (thickness, length, width) mild steel plates. The thickness corresponds to that used in nuclear containment units. The width is such that no side interactions will take place when waves are directed down the length of the test platform and the width-centered source and receiver are not more than 70.0" apart (see Figure 7). This means that waves can be propagated down and back the nearly the entire length of the plate before side interactions are part of the measured return. Thus, for tests where propagation distances are less than 70 inches, the plate width is rendered effectively infinite.

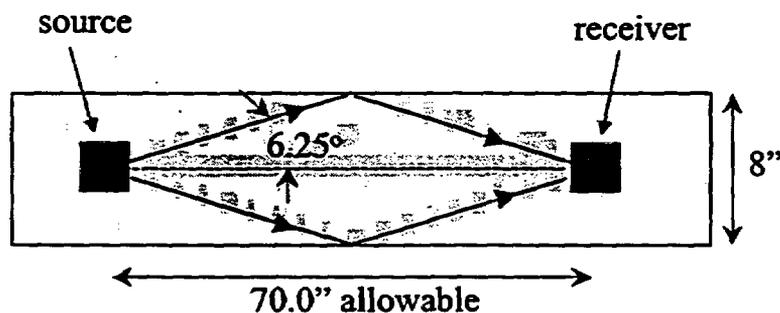


Figure 7 - Required plate width for the elimination of side interactions as competing signals.

In the following experimental procedures, all measured "degradations" are actually uniform cross sectional slots cut across the plates' width. Transducers were typically oriented perpendicular to the slots and because side interactions are intentionally avoided, the degradations are rendered effectively two-dimensional. The effect of a curved edge (i.e., not straight slots) on reflected returns is also an important recommended study for future research.

3.0 Experimental Procedures and Results

3.1.1 System Performance and Operability - Performance Check

As a first step in verifying the operability of the equipment, a conventional ultrasonic thickness test was performed. Figure 8 shows a sketch of the experiment and the measured return. A transducer was placed in direct contact (i.e., no coupling wedge) with the longest edge of a 8" x 1" x 38" steel plate, with the idea of injecting compressional waves injected across the plate's width. When operating in a monostatic configuration, the transducer acts as a receiver while it acts as a source, meaning that the input, free-vibration ringdown of the piezo element, and wedge reverberant field are all measured. This is shown as the clipped portion of the signal before 40 microseconds (receiver gain was set to record the reflected signal from the plate's opposite edge, which clearly displayed significantly lower levels than those signals listed above). The first returns after the source has died down are from the direct path to the opposite edge of the plate and they occur at 75 microseconds. When half of this delay (time of flight for one way travel) is multiplied by the assumed compressional wavespeed in steel (0.55 cm per microsecond) the 8-inch width is estimated to be 8.12 inches. Although the transducer used was not designed for this type of a test, the results point to a properly operating system.

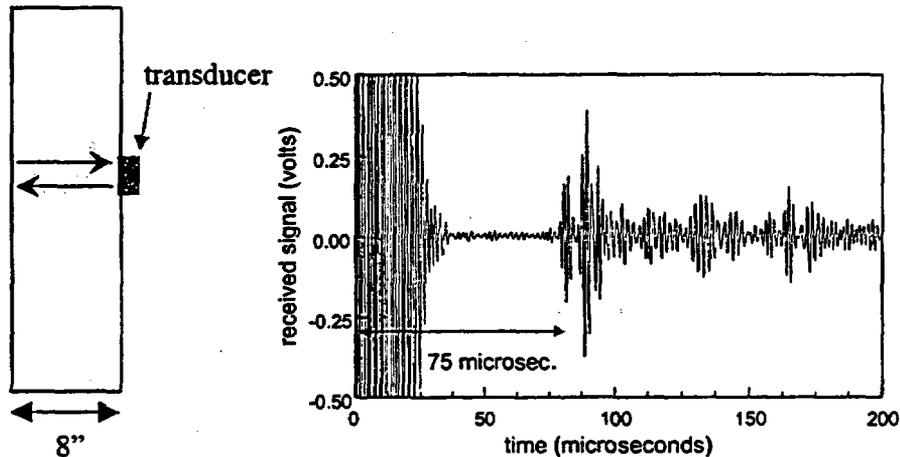


Figure 8 - Simulated UT test, setup and measured return.

3.1.2 System Performance and Operability - Dynamic Range

The system's input/output dynamic range was experimentally obtained by measuring the maximum signal that the system can inject and the minimum signal that the system can read. Two transducers (both attached to 45 degree coupling wedges) were placed adjacent to one another on a one-inch thick steel plate. With the receiver gain set to the allowed minimum (0 dB) and the input level set to the allowed maximum (100%), the signal transmitted from one transducer to the other was monitored. The receiver position was slowly varied away from the source until the received signal obtained its first maximum (i.e., the receiver fell on the first "hot spot"). This was assumed to be the

maximum signal level that could be injected. The source was then removed from the plate, the receiver gain maximized (70 dB), and one hundred averages of the no source signal assembled and averaged (to suppress uncorrelated sensor noise). This averaged signal was assumed to represent the minimum measurable signal.

Figure 9 shows the results. Plotted are the power spectral densities of the maximum and minimum signals. At 0.5MHz, where the source signal is concentrated, a 125dB difference can be seen. This difference represents the system's input/output dynamic range.

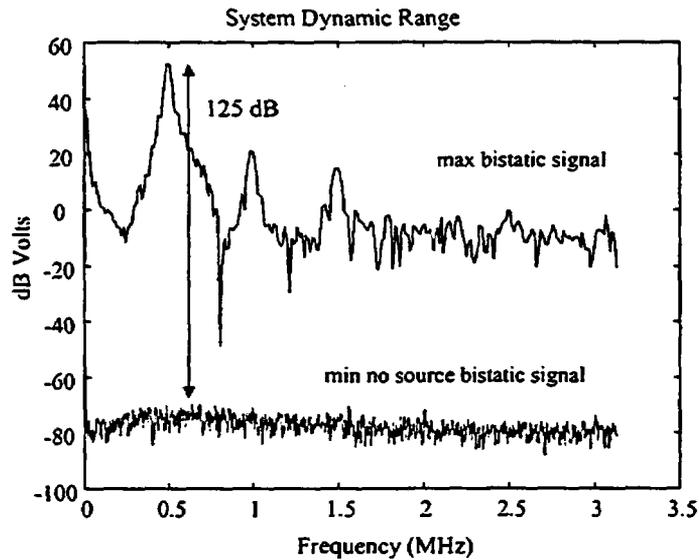


Figure 9 - Results for system dynamic range.

3.1.3 System Performance and Operability - Measurement Repeatability

In order to attribute differences in measured signals to physical effects, measurement repeatability was quantified. These tests were performed on vertically oriented plates to incorporate the problems associated with fixing the transducers in this orientation. The tests involved measuring the transmitted signal from one transducer to another for several different relative separation distances. Transmitted signals were recorded over a three-day period on 3 plates, for a total ensemble of nine measured signals at each separation distance. Of those nine signals, the ratios of the maximum to minimum received signal levels are plotted in Figure 10 for 5 different source/receiver separation distances. The plot shows a maximum variability of 2dB, thus allowing measured differences greater than 2dB to be attributable to effects outside of pure chance.

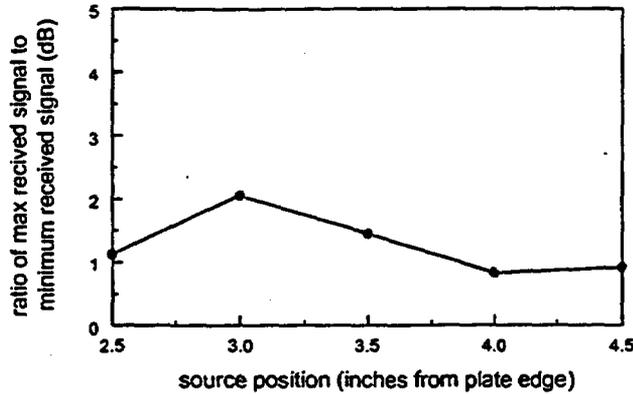


Figure 10 - Maximum variation in transmitted signal levels for various source-receiver separation distances.

3.2 Source and Waveguide Characterizations

In order to accurately image the location and shape of a degradation, the distortion caused by waveguide propagation of unfocused incident waves must be considered. For these purposes, numerical modeling will serve as an invaluable tool. However, the assumptions that must be made by those models will be critical. Therefore, as a basis for refining the numerical modeling assumptions, the waveguide effects on signal propagation were addressed in a series of tests.

Using one transducer as a source and a second as a roving receiver (with both transducers fixed to a 45° coupling wedge), a bistatic array was simulated. By doing so, the incident wave interactions with the plate's edges could effectively be monitored as it propagated down the plate. Fig 11 illustrates the experimental setup.

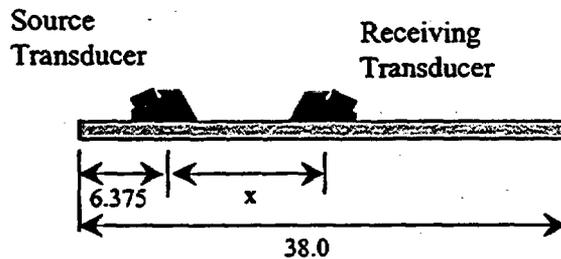


Figure 11 - Test setup for monitoring forward travelling waves in a free plate.

Figure 12 shows a plot of the measured signal for the forward propagating wave. The bottom axis represents time and the left-hand axis represents receiver distance from the source. Image brightness is proportional to the envelope of the received signal level. Note that the received signal takes longer to reach positions that are located further from the source.

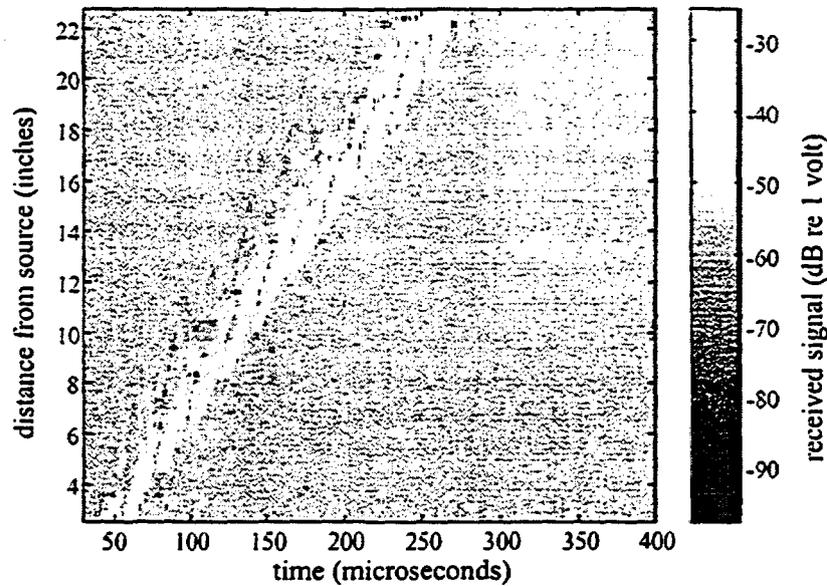


Figure 12 - Combined envelopes of received signals at several locations relative to a fixed source (forward travelling waves).

Several features of this plot can be brought to point:

- "Hot spots", seen as bright features, correspond to structural wave interactions with the measured surface. These features appear with regular spatial periodicity corresponding to the skip length of a 45-degree incident signal.
- The magnitudes of the hot spots diminish as the receiver is moved further from the source. The rate at which they diminish was verified to fit cylindrical spreading predictions, implying that the edge reflections are "specular", like a pool ball bouncing off the edge of a pool table. Moreover, because there is no strong evidence of scattering to non-specular waves due to surface imperfections, it can be concluded that the measurement noise floor due to surface imperfections will be low.
- A straight line can be fitted through the centers of the slashes, the slope of which corresponds to the global speed at which the energy travels down the plate. This global velocity is given by:

$$v_{global} = v_{shear} \cos(\theta_{incident})$$

This expression provides a convenient verification of the injected incident angle.

3.2.1 Assessment of Signal Loss Components - Degradation Shape

Figure 13 shows a schematic sketch of the test setup used to study the returned signal levels from degradations of various shapes. As noted previously, these

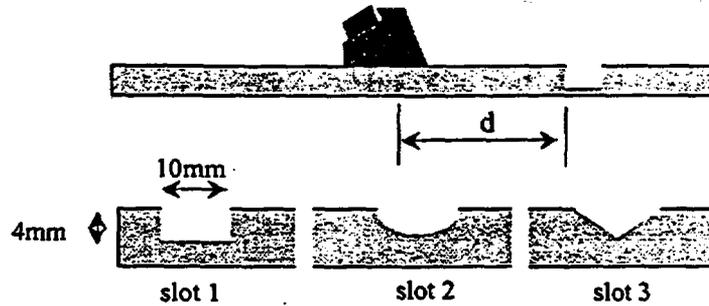


Figure 13 - Test setup for degradation shape study.

degradations are actually straight slots cut across the plates' 8 inch width. All of the degradations in this test were 4mm (0.158in) deep and 10mm (0.394in) in width. Monostatic returns were measured at several source locations relative to the slots' leading edges. Figure 14 shows a sample return signal. The reflected signal "level" is defined as the maximum value in the signal packet that is assumed, by simple time gating procedures, to have emanated from the slot.

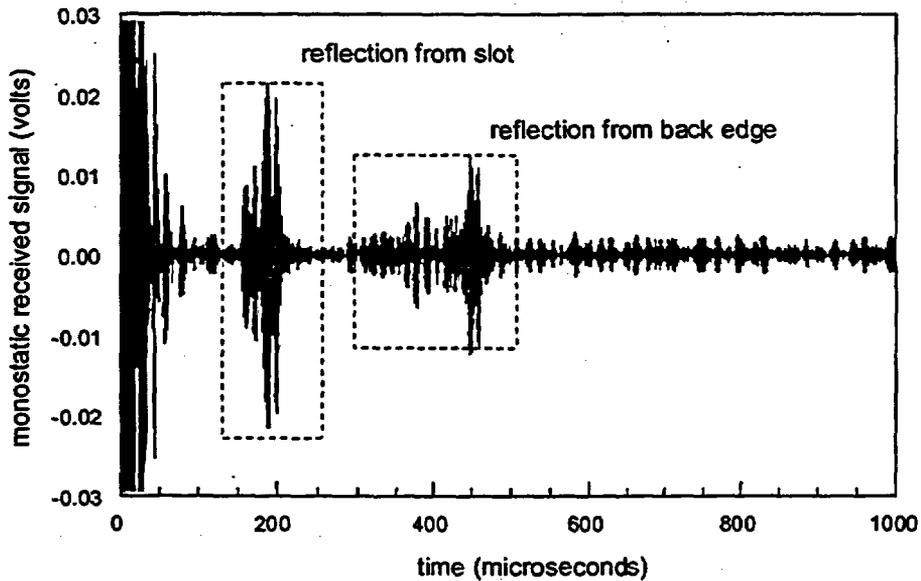


Figure 14 - Sample reflected return from rectangular slot (used to define "signal level")

Figure 15 shows the differences in reflected signal at several source locations relative to the slot. For the 45° wedge, the returns from the rectangular slot are, averaged over source location, around 1dB higher than those from the rounded slot and 4dB higher than those from the "v" shaped slot (see Figure 15a). For the 70° wedge, the differences are 5 and 9dB and are shown in Figure 15b.

These results include the effects of geometric spreading, which are greater for the 70° wedge than for the 45° wedge because the path to and from the slot is longer for the

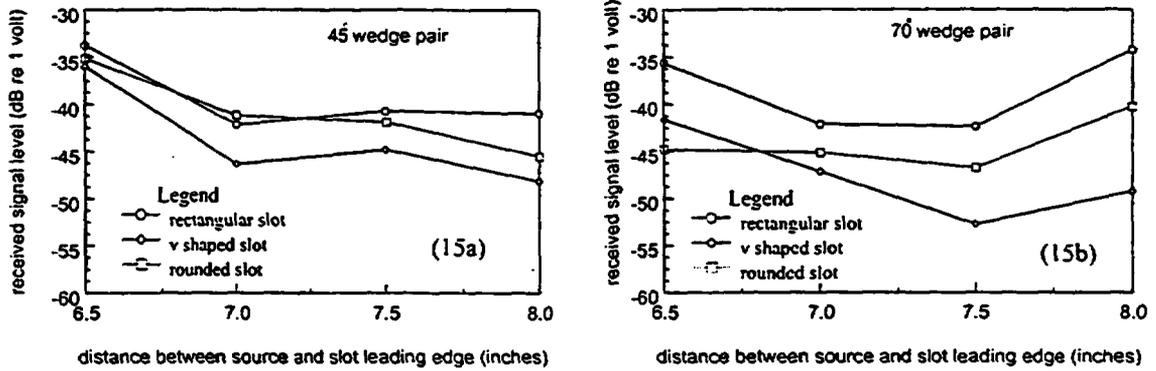


Figure 15 - Results from degradation shape study.

deeper angle. Additionally, the waveguide tends to compound the complexity of the reflected signal, because the incident field actually approaches the slot from a wide array of angles, and all reflected waves with any propagation component back in the direction of the source is measured as a reflected signal. A more thorough analysis of the measured differences would incorporate advanced acoustic scattering theories that are beyond the scope of this study.

Figure 16a shows examples the time windowed returns reflected from the three different slots. A comparison of these returns reveals that the arrival times of the reflected signals are very nearly equal. Figure 16b shows the ratio of spectral magnitudes for the rounded to the rectangular slot returns. It is in the frequency domain that the majority of scatterer characterization schemes will be implemented. Such characterization would require significant differences in the returns from different shape degradations. Although no attempt is being made here to perform frequency domain analysis, it appears as though there are significant differences not only in the level of the reflected signal, but also in the frequency character of the reflected signals.

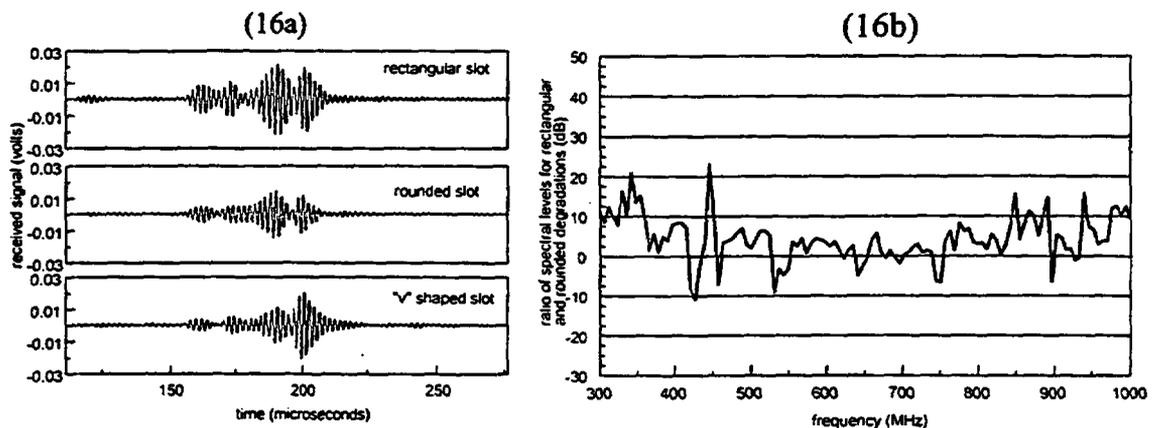


Figure 16 - Differences in reflected signal character for different degradation shapes. 16a shows windowed selections of the return signals reflected from the three different slots. 16b shows the ratio of spectral levels for the windowed return from the rectangular slot to the rounded slot.

Degradation depth was studied using the same methodology used to study degradation shape. Fig. 17 shows a sketch of the slots that were used in this test. All slots were rectangular in shape and 10mm in width. The slot depths were 4, 8 and 12mm. Reflected signal levels were determined in the same manner used in the shape study.

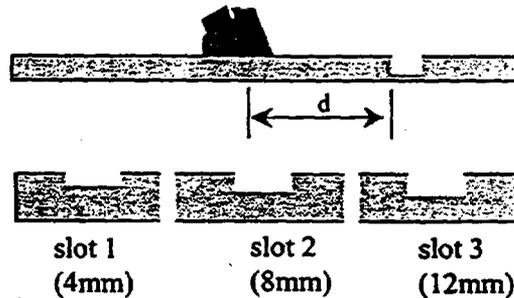


Figure 17 - Test setup for degradation depth study.

Figure 18a shows the differences in reflected signal at several source locations relative to the slot for both a 45° and 70° wedge pair. For the 45° set, the returns from the 4mm deep slot are, averaged over source location, around 3dB lower than those from the 8mm deep slot and about 6dB lower than those from the 12mm deep slot. This trend is as expected as deeper slots project a greater area of acoustic impedance.

Figure 18b shows the reflected signal levels from the three different slots using a 70° wedge set. The resultant trend for this case is not as expected. For some source - receiver separation distances the reflected returns from the shallowest slot are actually greater than those from the deepest slot. The trend was qualitatively observed over a broad range of separation distances, eliminating skipping effects as a possible explanation. The explanation for these unexpected results may, again, only be realized after incorporating advanced scattering theories.

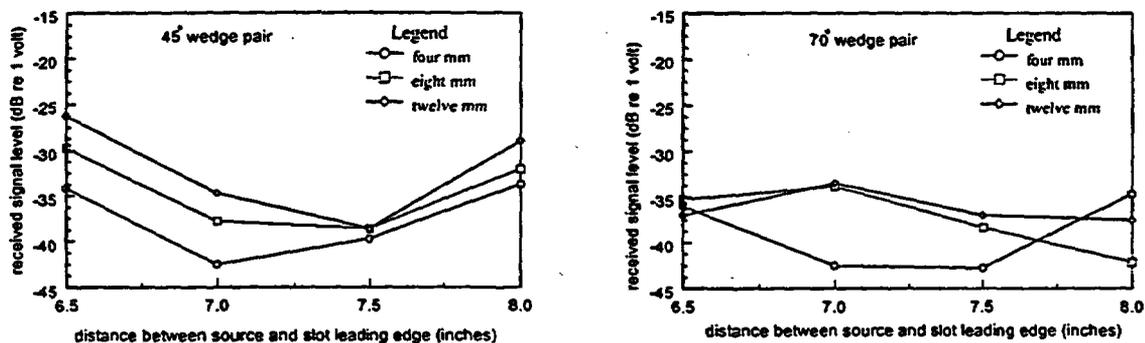


Figure 18 - Results from degradation depth study.

3.2.2 - Assessment of Signal Loss Components - Concrete Effects

The final bank of tests involved measuring the effects that concrete has on waves travelling in an embedded plate. Figure 19 shows a photograph of a wooden test base that was built to allow the midsection of three plates to be embedded in a concrete bath. The signal transmitted from one end of the plate to the other was measured before and after filling the molds with concrete. To determine the effect that bond quality has on induced losses, one of the plates was wrapped with a single layer of 4mil plastic sheet. The other two plates were tested under identical untreated conditions in order to quantify, albeit sparsely, the concrete effect's repeatability. These plates are referred to either as "coated", i.e., wrapped with a plastic sheet, or untreated, i.e., no plastic sheet.



Figure 19 - Test base for concrete effects test.

Figure 20 shows a simplified schematic of the test setup for a single plate. The source transducer location was varied from 2.5 inches to 4.5 inches in 0.5-inch steps in order to

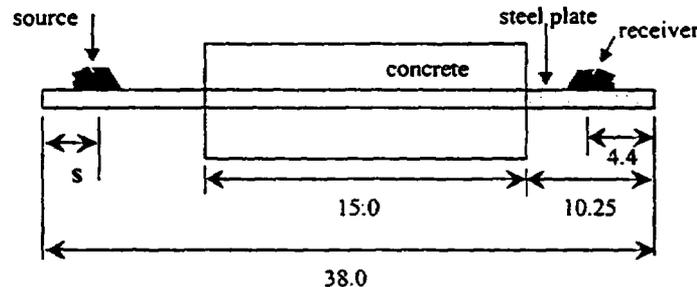


Figure 20 - Test setup for concrete effects test. All units are in inches.

sufficiently sample one half of a skip length for the 70 degree wedge. The wave propagated through the embedding region was monitored by a fixed position receiving transducer. The received signal level is, as before, defined as the maximum value in the incident signal packet reached by the receiver.

Figure 21a shows the signal level received for three free standing (i.e., no embedding concrete) plates, for several source locations. As expected, there is only

minimal variation from plate to plate and the plastic wrap has no noticeable effect. Also note that the received levels are about 100dB above the measurable floor, with all losses being attributable to geometric spreading. Figure 21b shows the received levels after embedding the midsections of the plates in concrete. The two uncoated plates display a significant loss in signal level (30dB, or 1.6dB per centimeter of two-way travel) while the coated plate level remains relatively high, incurring virtually no losses. Note that the results for the untreated plates are very similar, with the differences being attributable to a combination of measurement variation and concrete bond variation. Upon removing the plates from the test base, the concrete surrounding the wrapped plate adhered well enough to support its own weight. Thus, it appears as though the micro-character of the concrete-to-steel bond plays a critical role in determining the proportion of energy lost from waves travelling in the steel.

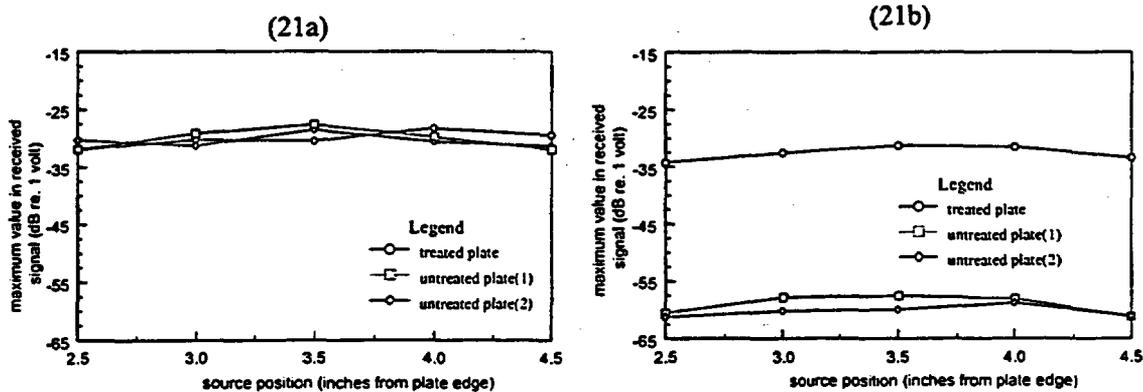


Figure 21 - Transmitted signal level for three plates (one coated, two uncoated) free-standing (21a) and partially embedded in concrete (21b)

Figure 22 shows the results for the transmitted signal levels through the coated plate partially embedded in wet and cured concrete. Note that the transmitted levels are 3-4dB lower for the wet concrete case.

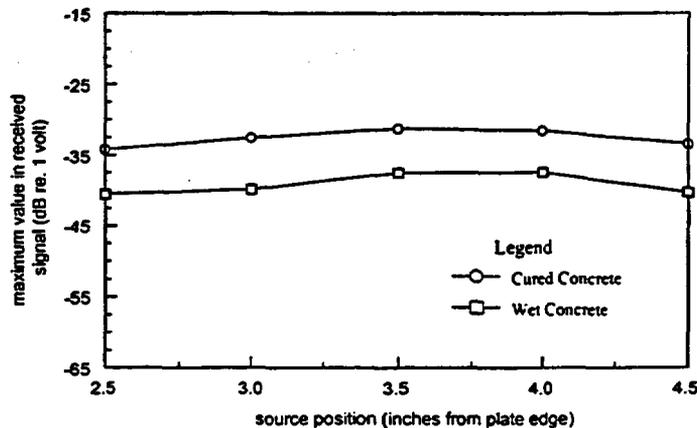


Figure 22 - Transmitted signal levels for the coated plate partially embedded in wet and cured concrete.

Figure 23 shows the received signal levels using different wedges tested on a single untreated plate after it was embedded in concrete. The 70-degree wedge shows

transmitted signal levels that are roughly 4-5 dB higher than those received using the 45-degree wedge. Assuming that both wedges couple equally well to the test structure, this result is to be expected because the incident wave injected by the 70-degree wedge experiences fewer interactions with the concrete-steel interface due to the longer skip distances.

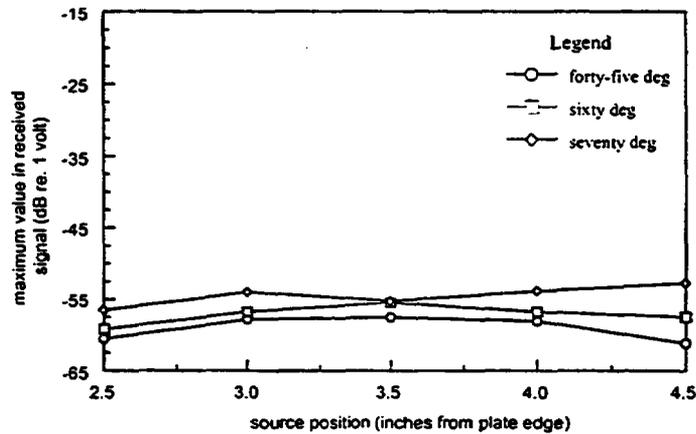


Figure 23 - Transmitted signal levels for embedded plate using different wedge angles.

3.4 Practical Detection Issues - Masking Signals

The objectives of these experiments were to determine the system loss components in order to provide a total loss estimate and verify that the total loss does not exceed the system's dynamic range. However, these procedures do not address the problem of competing signals. To gain a feel for the measurement floor that is dictated by the presence of competing signals, the signals reflected from a rectangular slot (4mm deep) located five inches below the air-concrete interface were measured (see Figure 24). The results were compared to those obtained with no concrete present. This test represents a scenario that one might encounter in the field and thus is an important gauge for the technique's feasibility.

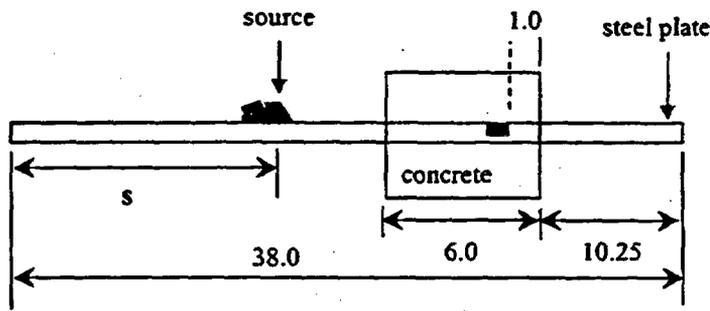


Figure 24 - Test setup for simulated field scenario.

Figure 25 shows the return from the slot after concrete has been poured. The signal level in the reflected signal "packet" is down about 20dB (linear factor of 10) relative to the reflected return from the slot without concrete present. The return from the slot is still evident, but the competing return signals (occurring between 50 and 130 microseconds) that would normally be considered secondary, are now only 6-10dB down (linear factor of about 2-3) from the peak in the reflected packet emanating from the slot.

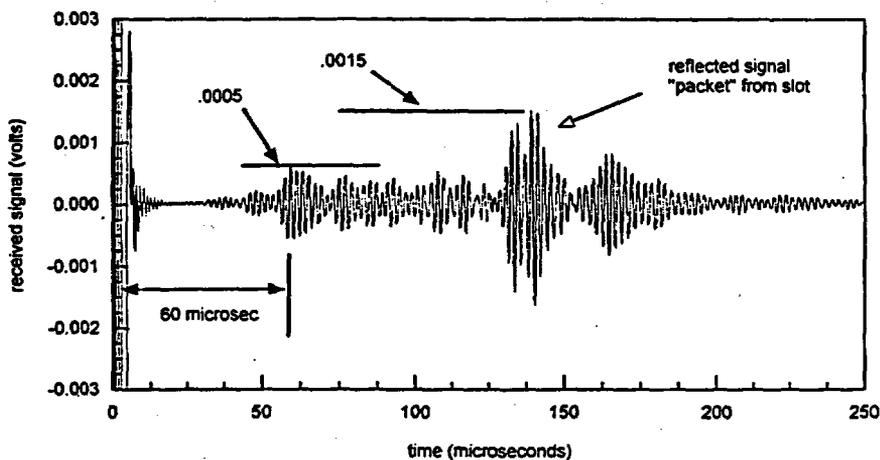


Figure 25 - Reflected signal from an embedded rectangular slot.

It is interesting to note that the reflected signal builds somewhat sharply at around 60 microseconds. This time delay corresponds very nearly to the time it would theoretically take a compressional wave propagating directly down the plate's axis to and from the slot's leading edge. Thus, it appears as though an incident side lobe has coupled to a compressional wave in the steel. Therefore, these signals could be considered a localizing aid as opposed to competitors. Consequently, no quantitative conclusions can be drawn from this experiment concerning the relative levels of competing signals. The important point is that if a competing signal can be identified, then it can in all likelihood be removed from the return signal during post-processing. Implementing advanced signal processing algorithms to exploit and/or discriminate phenomena of this type is an important recommendation for future work.

4.0 Conclusions

Corrosive thinning and pitting of the pressure vessel walls in embedded regions of steel nuclear containment units currently cannot be detected without employing expensive, dangerous and potentially inconclusive concrete chipping methods. It has been proposed that high frequency acoustic imaging may serve as an alternate solution to detecting these degradations.

A numerical study has already been carried out to address preliminary feasibility issues. As a basis for improving these numerical models, and also to continue the study in a more practical setting, a series of controlled laboratory experiments were executed. The experiments were designed to assess the performance of a commercially available fully integrated angle beam inspection system, and to further investigate the underlying physics that govern the use of angle beam inspections. To pursue these goals, an angle beam inspection system was obtained from Matec Instruments, Inc.

The experimental results are as follows:

- The measurement system displayed an input/output dynamic range of 125dB. Therefore, in the absence of competing signals, 105dB of losses can be incurred while still maintaining a 20dB signal-to-noise ratio. The system tested is of moderate quality, and the dynamic range could in theory be increased by as much as 35dB by selecting a more powerful pulser and a higher precision acquisition board. The system displayed measurement variations on the order of 1-2dB, which implies that characterization of degradation dimensions will be subject to a 25% margin of error and that measured differences greater than 2dB can be attributed to physical effects. The source was verified to inject waves at refracted angles very near those specified by the manufacturer.
- The mild steel plates that were used in the experiments propagated signals as if they were effectively free of surface imperfections. Under similar conditions, the surface generated noise floor and the false-positive alarms expected from surface flaws should be minimal.
- The reflection coefficient from a 4mm deep, 10mm wide two-dimensional degradation, measured using a 45-degree wedge, was shown to be about -24dB at 0.5 MHz. With the same wedge, rounded and "v" shaped degradations of similar depth and width showed returns on average 1dB and 4dB, respectively, lower than those returned from the 4mm deep rectangular degradation. For the same wedge, returns from a 8mm deep rectangular degradation showed returns 3dB higher than those from the 4mm deep degradation and a 12mm deep slot showed returns 6dB higher. Using a 70-degree wedge, rounded and "v" shaped degradations of similar depth and width showed returns on average 5dB and 9dB respectively, lower than those returned from the rectangular degradation. The relation between reflected signal level and degradation depth using the 70-degree wedge is unclear. These results provide a preliminary basis for estimating reflected signal levels from a vast array of

two-dimensional degradations.

- When plates are embedded in concrete, an additional 1.6dB of signal loss is incurred for each centimeter of two-way travel using 45-degree wedges, and 1.4dB using 70-degree wedges. Plates that have a plastic wrap between themselves and the embedding concrete show virtually no additional losses compared with free plate signal losses. Therefore, in regions where corrosion is suspected, it is likely that the concrete-to-steel bond is compromised, and the effect of the concrete may be minimal. However, the presence of water in the concrete may significantly increase the signal loss that is induced.
- The results from the signal loss components experiments can be combined to provide a basis for estimating the total loss induced on an incident signal for many scenarios. For example, a 4mm deep rounded degradation, located 30cm below the air/concrete interface should display reflected signals 76-78dB down from the incident signals measured at the drive point. This does not include geometric spreading, which will add an additional 7dB of loss (based on 4dB of spreading loss at 16cm, as noted from Figure 15a, which shows return levels of around -35dB for the rounded degradation at a distance of 6.5 inches or 16cm, 25dB of which are due to the reflective character of the degradation and based on a maximum of -6dB). In the absence of competing signals, a degradation of this type should be detectable because roughly 40dB of signal-to-noise ratio remains (based on the system's input/output dynamic range).
- Rectangular degradations 4mm in depth located 5 inches below the air/concrete interface can be detected without addition signal post-processing. The competing signal environment, which tends to mask returns from actual degradations, can be accounted for using basic wave propagation analyses. Techniques to either utilize or discriminate against these secondary mechanisms must be implemented for the technique to demonstrate success in more challenging scenarios.

5.0 Recommendations

The results from this project can be used to determine, on a preliminary basis, the feasibility of employing ultrasonic imaging in a vast array of degradation scenarios. It appears as though moderately sized corrosive degradations (4mm and greater depth with fairly abrupt edges) can be detected to distances of around 30cm below the air-concrete interface. However, the study does not address several factors that could stand in the way of eventually employing the technique.

Recommendations for future work fall into two main categories: continuation of the feasibility study and the development of more sophisticated detection, localization and degradation characterization solutions.

Additional feasibility issues include:

- Three-dimensional effects, such as plate curvature, degradation curvature and the mixing of return signals from several closely spaced degradations.
- Effects of structural discontinuities in the pressure vessel, such as periodically located concrete anchors.
- Studying in significantly more detail the actual shapes that corrosion typically takes on.
- Studying the bond quality between concrete and steel in areas of corrosion.

In order to develop more sophisticated imaging algorithms, the effective spatial and temporal filters that an incident and reflected signal pass through must be well known. Therefore, a thorough study of the acoustic properties of transducers and wedges must be carried out. In addition, the ability to remove propagation path information from the reflected signal must be studied. This will be difficult in embedded scenarios without knowing the concrete-steel bond quality, which was shown to play a major role in determining total induced loss. Once both of those studies are successfully completed, it will then be possible to carry out a study on implementing inverse scattering techniques for the purpose of characterizing degradation dimensions.

Several procedural recommendations can be made based on the practical knowledge gained in this study:

- The transducers used act as relatively narrowband mechanical filters (in relation to conventional, highly damped transducers) and are tuned to 0.5 MHz nominally. Frequency concentrated tone burst waveforms were used to minimize the energy lost outside of the effective transducer filter and therefore to maximize the injected power. In retrospect, the added problems of a relatively lightly damped piezo crystal (which had a significant ringdown period) and a relatively long input signal (required for frequency focusing) do not justify their selection over a highly damped crystal excited with short pulses.
- It is imperative that gel couplant completely fills the space between the transducer-wedge interface and the wedge-test structure interface. This becomes increasingly difficult when wedges are coupled to vertical structures.

Acknowledgements

The authors wish to acknowledge Robert Sammataro for his insight into regulatory issues and Joe Edwards for his help in preparing the manuscript. Oak Ridge National Labs and the Nuclear Regulatory Commission have sponsored this work.

References

- 1 J. Bondaryk, C. Corrado and V. Godino, Feasibility of High Frequency Acoustic Imaging for Inspection of Containments, NUREG/CR-6614 (ORNL/SUB/97-SX754V), Martin Marietta Energy Systems, Inc., Oak Ridge National Laboratory, Oak Ridge, Tennessee, August 1998.
- 2 M.C. Junger and D. Feit, Sound, Structures and Their Interactions, 2nd ed. (MIT Press, 1986, 2nd printing ASA/AIP, Woodbury, NY 1993).
- 3 L. Kinsler and A. Frey, Fundamentals of Acoustics, 2nd ed., Wiley and Sons, NY, 1962.

OVERVIEW OF THE OECD - HALDEN REACTOR PROJECT

Carlo Vitanza

OECD Halden Reactor Project

P.O. Box 173, N-1751 Halden, Norway

Tel: + 47 69212200; Fax: + 47 69212201

ABSTRACT

The OECD Halden Reactor Project is an international network dedicated to enhanced safety and reliability of nuclear power plants. The Project operates under the auspices of the OECD Nuclear Energy Agency and aims at addressing and resolving issues relevant to safety as they emerge in the nuclear community. This paper gives a concise presentation of the Project goals and of its technical infrastructure. The paper contains also a brief overview of results from the ongoing programme and of the main issues contemplated for the next three-year programme period (year 2000 - 2002).

1. INTRODUCTION

Safe and reliable operation of nuclear power plants benefit from R&D advances and related technical solutions. The OECD Halden Reactor Project is a leader in these advances with programmes devised to provide answers in a direct and effective manner. The Project's strong international profile and solid technical basis represent an asset for the nuclear community at a time in which maintaining centres of expertise at accessible cost becomes increasingly important.

The Halden Project is a joint undertaking of national organisations in 20 countries sponsoring a jointly financed programme under the auspices of the OECD - Nuclear Energy Agency. The programme is to generate key information for safety and licensing assessments and aim at providing:

- Basic data on how the fuel performs in commercial reactors, both at normal operation and transient conditions, with emphasis on extended fuel utilisation.
- Knowledge of plant materials behaviour under the combined deteriorating effects of water chemistry and nuclear environment.
- Advances in computerised surveillance systems, human factors and man-machine interaction in support of upgraded control rooms.

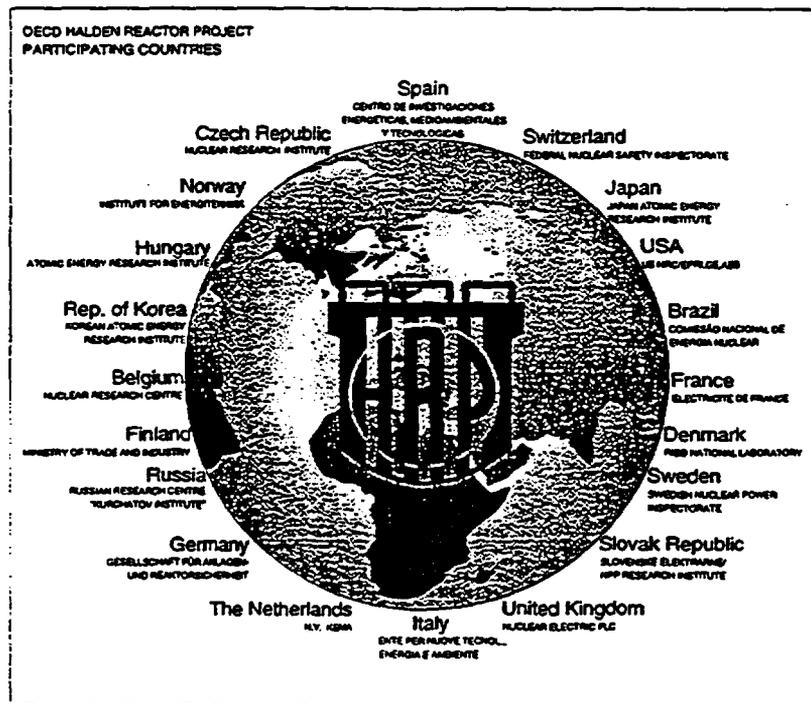


Fig. 1. The Halden Project is an international network with 20 member countries. The participants represent a complete cross section of the nuclear community, including regulatory bodies, vendors, utilities and R&D centres.

In addition to the joint programme work, a number of organisations in the participating countries execute their own development work in collaboration with the Project. These bilateral arrangements constitute an important complement to the joint programme and normally address issues of commercial interest to a participant organisation or group of organisations.

The programme results are systematically reported in Halden Work Reports and in Enlarged meetings organised by the Project. Participants' bilateral activities are also presented at these meetings. Special workshops with participation of experts are frequently arranged for in-depth assessments of specific issues, especially when new programme issues are to be established.

The joint programme is renewed every third year. The programme renewal involves extensive reviews and discussions with Project participants on priorities, programme issues to be addressed and technical means to achieve the programme objectives. The large circle of participants, with the consequent cost-sharing among many parties, has enabled to utilise the overall infrastructure to the maximum possible extent. The Halden Project is committed to continue this endeavour by responding efficiently to technical requirements emerging in the nuclear community, by maintaining its facilities in good order and by continuing to adhere to a highly competitive cost structure. Norway, host country, has always been strongly supportive of the Halden Project and is expected to do so in the future. The Norwegian contribution covers 30% of the joint programme funding.

2. FUEL AND MATERIALS PROGRAMME

2.1 Key Facilities

The main tool for the fuel and material work is the Halden Boiling Water Reactor (HBWR), with its range of experimental capabilities. The license for the reactor is to be renewed in 1999, and as on previous occasions, the Norwegian Institute for Energy Technology, which operates the Halden Project facilities, has applied for a renewal for a period of ten years. The license application is based on an extensive review of the Safety Report for the HBWR, particularly regarding the conditions of the vessel. Data so far show that fluence-induced damage progresses at a low rate and that current criteria result in a vessel projected lifetime well beyond year 2020.

Substantial development has taken place at Halden, aimed at providing a flexible facility where a variety of experimental needs can be accommodated. When specific coolant conditions are required, such as for cladding and structural materials studies, water loops are available. The loops can be operated in different thermal-hydraulic and water chemistry conditions, covering a range of BWR and PWR requirements.

The distinctive speciality of the HBWR fuel and material experiments resides in the ability to perform high quality in-reactor measurements, which provide unique and well characterised data during operation; that is, while mechanisms are acting. The Project experimental programmes are centred around this capability and make use of it to the maximum possible extent.

This capability has in recent years been extended such that commercial fuels can be efficiently tested at Halden. Fuel rods extracted from commercial reactors can be segmented and re-fabricated into rodlets suitable to further specialised testing at Halden. For this purpose the fuel segments are also retrofitted with the instruments required for the tests.

Similarly, structural materials extracted from LWR cores can be machined, fatigue pre-cracked if necessary and suitably instrumented for the Irradiation Assisted Stress Corrosion Cracking (IASCC).

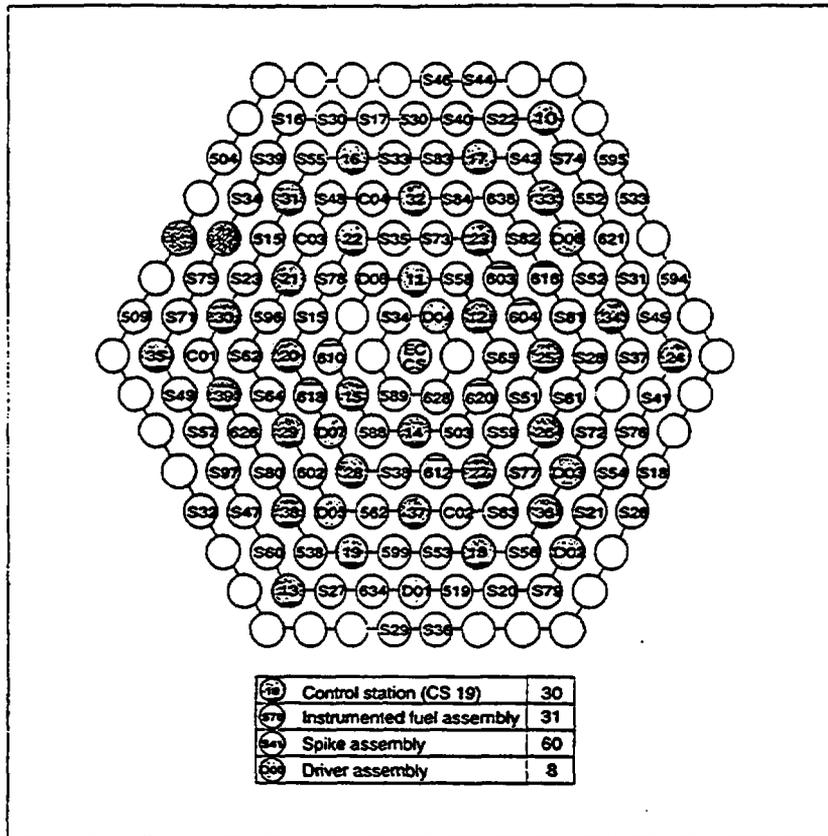


Fig. 2. Cross section of the Halden reactor core. This core configuration refers to February 1998, with totally 31 test rigs and 68 driver fuel assemblies. About half of the test rigs are operated in LWR loops.

This technique is important in that it provide very representative materials already irradiated to doses typical of "aged" plants.

The experimental work is supported by the hot cells for re-fabrication and post-irradiation examinations, by workshops, electronics and chemistry laboratories and by a computerised Data Bank.

2.2 Ongoing Programme (Time Period 1997 - 1999)

The joint programme work in the fuel and material area focuses on the following main points:

- Characterisation of UO_2 fuel properties at high burn-up, typically up to 60 MWd/kg. The fuel thermal conductivity degradation is assessed by means of direct on-line fuel temperature measurements. The threshold for fission gas release and the fuel swelling are addressed in dedicated instrumented tests.

These investigations make use of both test fuel and commercial fuel rods. Four test rigs are dedicated to this objective.

- Extension of the fuel data base to include gadolinia fuel and MOX fuel. Substantial effort has been made to acquire representative fuels and produce new test rigs, especially for MOX characterisation. Both test fuel and commercial fuel are used in these tests. Two rigs are dedicated to gadolinia fuel and three to MOX fuel irradiations. The rods are instrumented. It should also be noted that one test rig is dedicated to the characterisation of commercial VVER fuel.
- Consequences of power and coolant transients to the fuel integrity. Power ramps (one test rig, reloaded) and short term dryout tests (one test rig, reloaded) have been carried out on pre-irradiated commercial fuel. The dryout tests have resulted in cladding temperatures ranging from -500 to -1000°C for -30 to 50 seconds. These tests are now completed. Preparation of a new test series addressing LOCA transients and the mechanistic understanding of RIA transients are underway.
- Consequence of increasing rod pressure at high burn-up. This study involves two rigs, one for cladding creep-out measurements in stress reversal conditions, the other for determining the pressure limit for cladding lift-off onset. Re-fabrication, instrumented commercial fuel rodlets are used for these tests.
- Corrosion of hydriding of modern cladding alloys at high burn-up. This test has started recently and aims at comparing the corrosion behaviour of a variety of alloys up to 50 MWd/kg . Another test rig has been used to determine mechanisms leading to secondary failures following (fuel and) cladding ID oxidation and hydriding.
- Irradiation assisted crack growth (IASCC) of as-fabricated sensitised and pre-irradiated stainless steel materials in normal BWR water chemistry and in hydrogen water chemistry. The crack growth is monitored on-line by means of in-reactor potential drop measurements. Pre-irradiated material is retrieved from commercial reactors, and then machined and instrumented at Halden. Two test rigs have been used in the ongoing programme period, a third rig is under preparation.

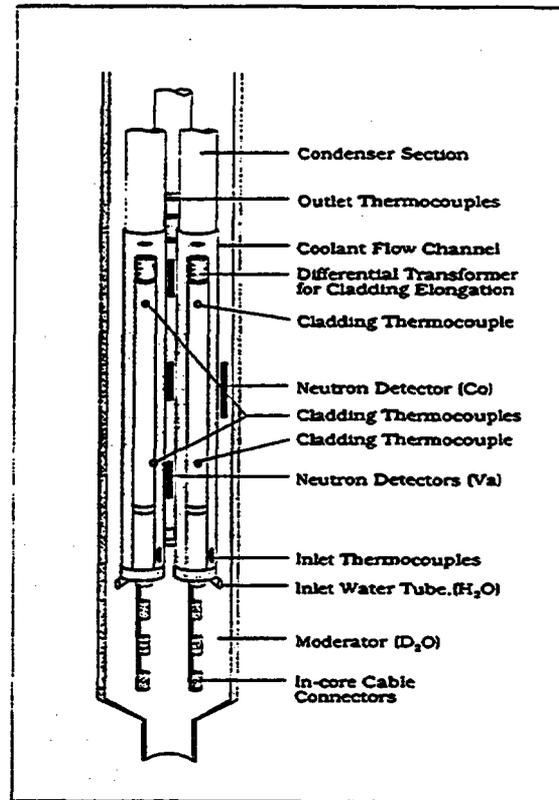


Fig.3. Test rig for determining the consequence of short term dryout. Three pre-irradiated, commercial fuel rodlets, after service in a BWR, were re-fabricated, instrumented and loaded into this rig. Each subchannel which contained one fuel rod, could be individually operated at reduced coolant flow conditions, producing dryout in the upper portion of the fuel rod.

- Initiation of IASCC in sensitised stainless steel, aiming at determining the onset of IASCC as function of stress and fast neutron fluence. One rig is being used for this investigation.
- Effect of alloy composition and fast neutron fluence on the susceptibility to IASCC. This test is carried out in collaboration with the USNRC/ANL and involves one irradiation rig. The specimens have been fabricated and are PIE-tested at ANL.

2.3 Boundary Conditions and Needs

In many countries utilities are faced with intense competition due to deregulation and in order to compete effectively, they are looking to improve operational economics and flexibility. At the same time, regulatory authorities have to verify that this is done without detriments to reactor safety.

Licensees are implementing or considering extended burn-up, longer fuel cycles, power upratings and load follow as means to reduce operational and fuel cycle costs. This exposes the fuel to increasing challenges, which has prompted the vendors to propose new fuel designs and new materials. There is also a strong push to use mixed oxide fuels in power reactors.

Regulatory bodies are faced with the need for qualified models and codes for safety case assessments in a variety of operational conditions, for many different types of fuel designs and at extended burn-up. This necessitates new and improved data on fuel properties and fuel behaviour under various normal, abnormal, and accident conditions.

At the same time, operational experience demonstrates that unforeseen anomalies can develop as demands on performance become more stringent. Localised corrosion and defected fuel degradation are potential utility concerns. Control rod sticking and anomalous axial power offsets have recently posed limitations on plant capacity factors and caused regulatory concern. Regulators will have to assess the consequences of these anomalies and determine effective surveillance practices, whilst the industry has to find valid technical remedies. Halden experiments can be of great value for addressing and resolving these issues and those likely to emerge in the future.

As the age of power plants increases, safety authorities will need materials property data relevant to in-reactor components at high irradiation doses, as they will form the basis for plant lifetime assessments. Utilities are introducing operational changes that can enhance the reliability of plant structural components - e.g., water chemistry modifications - and are adopting advanced materials where this can be done. Pressure vessel annealing is a possible option for mitigating the effects of radiation embrittle-

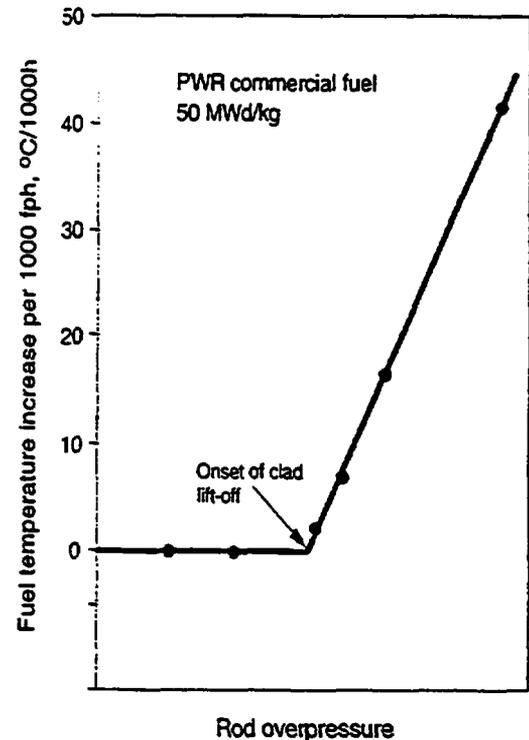


Fig. 4. Result of a test conducted on PWR fuel to determine the pressure limit before onset of cladding lift-off. It was shown that the fuel could withstand substantial overpressure (beyond PWR system pressure) without any sign of lift-off.

ment. Practical verifications and data will be needed to support lifetime predictions of existing and replacement materials as well as to validate measures intended for lifetime extensions.

2.4 Aims of the Programme for the Time Period Between Year 2000 and 2002

The experiments contemplated here aims at determining potential burn-up limiting phenomena. More emphasis will be placed on fuel swelling and pellet-cladding interaction. Tests will also address the gas mobility in high burn-up fuel rods, since this is believed to impact the response to LOCA (as ballooning is sustained by gas flow along the rod) and also to RIA. It is foreseen that cladding corrosion and hydriding will be more extensively addressed, since the status of the cladding at normal conditions may greatly affect the response in safety transients. The data will be used by Project participants as reference for the fuel codes assessments.

The proposed programme for the time period between year 2000 and 2002 focuses on the following main issues:

- *Fuel high burn-up capabilities in normal operating conditions*, aiming at providing fuel property data needed for design and licensing in the burn-up range 50 to 80 MWd/kg. In selected tests the burn-up will be pushed up to 100 MWd/kg. Both test fuel and re-fabricated commercial fuels will be used in the proposed investigations.
- *Fuel high burn-up capabilities in safety transients*, aiming at providing experimental complements to investigations conducted elsewhere on loss of coolant and reactivity transients. The LOCA tests are intended to address high burn-up, integral rod behaviour during transients and to complement separate effects investigation conducted, for instance, at Argonne National laboratory. The RIA investigation will instead focus on supporting the mechanistic understanding of the RIA transient, in particular the role of fuel swelling and fission gas release. Further tests on short-term dryout and new tests on power-coolant flow oscillations are also considered. The latter are intended to respond to regulatory priorities on Anticipated Transients Without Scram, where needs have been set forth for verification of the enthalpy criterion at high burn-up.
- *Fuel performance anomalies*, aiming at determining the cause for fuel anomalies to occur during service, as well as at identifying realistic design or operational remedies. These investigations are to be conducted in synergy with bilateral activities. The items under discussion include
 - Crud deposition as affected by water chemistry and heat rating.
 - Axial offset anomalies caused by local boron accumulation on the surface of PWR fuel rods.
 - Degradation of failed fuel resulting in large exposure of the fuel to the coolant and consequent increase of radiation level in the coolant.
 - Control rod sticking as result of axial growth of guide tubes during service.

It must again be clarified that the joint programme cannot address all these issues and that the Halden Project work scope must be put in the context of what is being done in other programmes.

- *Plant lifetime assessments*, aiming at generating validated data on stress corrosion cracking of reactor materials at representative stress conditions and radiation/water chemistry environment. The work initiated in previous programme periods on irradiation assisted stress corrosion cracking will be extended to include highly irradiated materials. The programme is intended to clarify the extent to which remedies introduced to alleviate the stress corrosion of in-reactor components remain applicable to components which have been in service for a long time. One focus will be on

representative BWR materials which have been retrieved from commercial reactors and on the use of these materials for in-core measurements of crack growth rates at given stress intensities. A second focus will be on stress corrosion studies under PWR conditions, as it is anticipated that cracking of in-reactor materials in PWRs may also become an issue of concern.

The embrittlement of reactor pressure vessel materials due to neutron irradiation is an important issue as nuclear plants age and is also addressed in the programme proposal. The Project intends to support collaborative programmes with participants in this area as needs arises, notably by utilising the Halden reactor as a source of neutrons under a wide variety of temperature and flux and fluence conditions.

3. MAN-MACHINE INTERACTION PROGRAMME

3.1 Key Facilities

The Virtual Reality (VR) centre currently established at Halden is a complement to HAMMLAB, providing the basis for new developments in control room design and engineering, particularly for control room upgrades. Applications are also envisaged in decommissioning, particularly in relation to design of special decommissioning tools, operational procedures, and training with maintenance procedures.

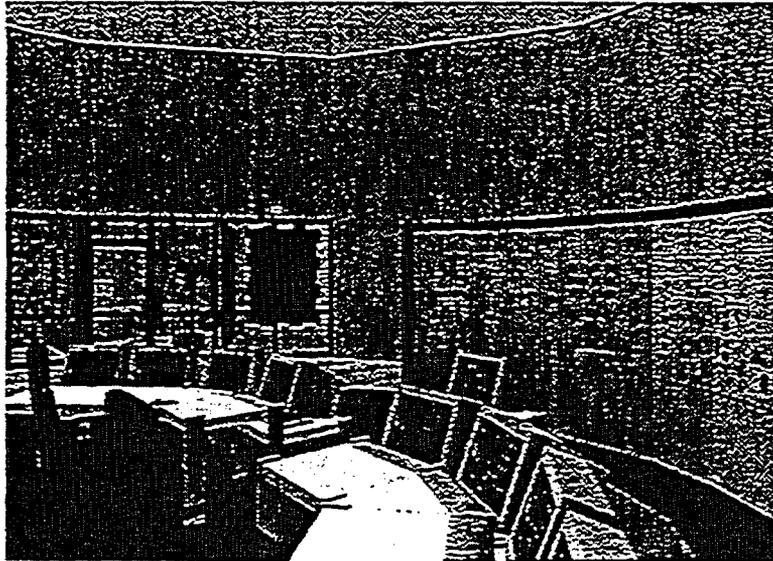


Fig. 5. Layout of a control room produced for a utility using the Halden Virtual Reality facility. This was part of the design for the modernisation of the control room in a nuclear power plant in Europe. The Halden VR facility is presently used mainly in control room engineering - in conjunction with the Halden human factor expertise. Expected applications include computer-based training and decommissioning.

The simulator-based Halden Man-Machine Laboratory (HAMMLAB) is the main vehicle for the human-machine systems research at the Project. A programme for expanding and upgrading the HAMMLAB is being undertaken to enable the facility to meet future requirements for human factors and control room research. One goal is to establish a flexible infrastructure in terms of hardware and software tools, including powerful, modern full-scope simulators for PWR, BWR and VVER systems. Thus, the simulators will be able to reproduce relevant power plant systems during normal, disturbed, and accident conditions. The built-in functionality of the HAMMLAB 2000 will facilitate the transfer of new results to commercial power plants.

The activities in the HAMMLAB and in the VR centre rely on the availability of simulators and of computerised operator support systems. Such tools should continuously be updated and utilised in control room applications by Project participants.

3.2 Ongoing Programme (Time Period 1997 - 1999)

- A series of pilot studies have been performed to find a reliable and valid methodology to investigate *human error in a control room setting*. These pilot studies investigated diagnostic strategies and styles that had been observed in earlier single operator and team based studies.
- One of the issues addressed by the first main *human error experiment* is the question of detection and recovery of erroneous actions in the situation where they actually take place. Since detection clearly is a prerequisite for recovery, the experiment considered how well people are able to detect the erroneous actions they make and how the level of detection depends on the circumstances or conditions, such as interface, team, workload, etc. The other main purpose was to develop a method for predicting performance failures, specifically the error modes that can be expected for a specific task.
- Systematic experiments in HAMMLAB put special demands on *operator performance measurements*. Performance scores have to be comparable across scenarios and sensitive to a wide range of operator competence levels. The measure should be reliable and robust, should account for team performance as well as single operator performance, and be efficient in use. It is furthermore desirable that the measure complies with established norms for human performance measurement regarding reliability, validity, sensitivity, non-intrusiveness, etc. The Project has therefore developed a method based on the prescription of optimal solutions to scenarios, based on discussion with process experts. For each scenario, the expert solutions are represented hierarchically in a diagrammatic form. Operator activities are classified and weighted according to their importance. During the experiment, the process expert registers operator activities in real time, concurrent with operator performance. After the study a performance index is calculated estimating the discrepancy between the expert analysis and operator solutions to the scenarios. In 1997 the Project analysed data collected during earlier alarm and human error experiments. The results show a good consistency among the activity types defined by the system, and a moderate relationship with plant performance.
- A method for analysing *plant performance measures* has been developed and applied to data collected from crews participating in experiments. The performance of the crew is compared to an optimal control model developed for the scenario and afterwards a data analysis technique is applied. The applicability of PPAS is promising, and it might be an important additional measure based on real plant measures to evaluate the quality of the crew performance.
- Project staff has designed a so-called integrated large overview display with the idea to support rapid assessment of the plant status and dynamics by a representation of the whole process. The display is

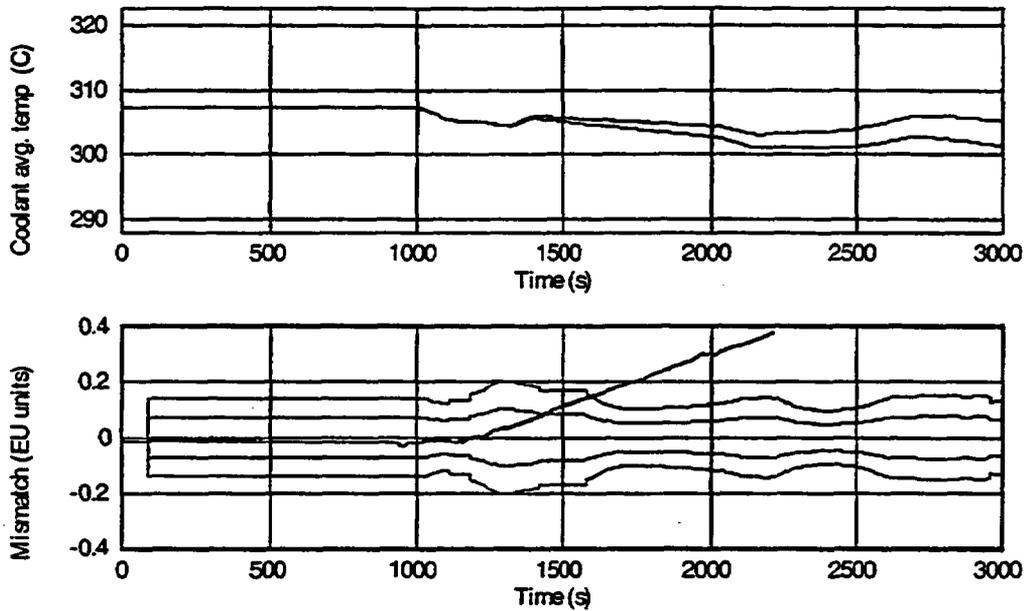


Fig. 6. Results from performance tests with the signal validation toolbox PEANO. The data were supplied by EDF/CEA, France. The system was able to track immediately a sensor failure, and, based on earlier "learning", reconstruct a best estimate value for the signal.

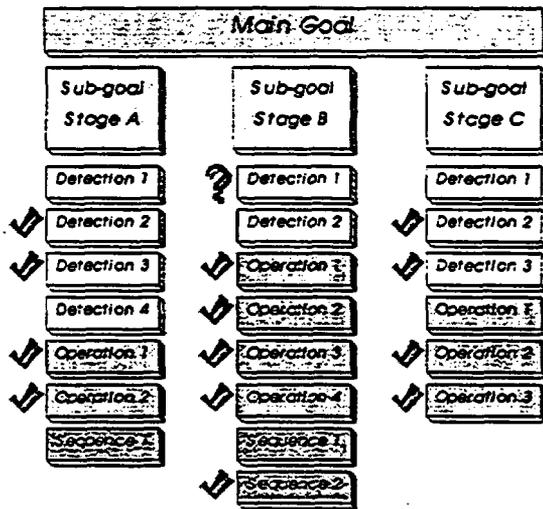


Fig. 7. Experiments in HAMMLAB put special demands on operator performance measurements. The Project has developed a method based on the prescription of optimal solutions to scenarios, based on discussion with process experts. For each scenario, the expert solutions are represented hierarchically in a diagrammatic form. Operator activities are classified and weighted according to their importance. During the experiment, the process expert registers operator activities in real time, concurrent with operator performance. After the study, a performance index is calculated.

shared by the control room staff and is a strong support for co-ordination. The layout and content of the display are context dependant to match the changing operators' needs and tasks. Centred around mimic diagrams, it combines different graphical features to support an efficient control of the complex process. Configurable elements and an original alarm presentation support clear and rapid identification of disturbances.

- The surveillance and control task of any industrial plant is based on readings of a set of sensors. It is essential that the output from these sensors are reliable since they provide the only objective information about the state of the process. The *signal validation task* confirms whether sensors are functioning properly. A method for transient and steady state on-line signal validation has been developed at the Project using artificial neural nets and fuzzy logic pattern recognition. The method has been successfully tested on simulated scenarios covering the whole range of PWR operational conditions. Data was provided by EDF, France. The neuro-fuzzy model has been implemented in a client/server software system under Windows NT. The system is called PEANO.
- Following the principles behind *formal software development*, the Halden Project has developed a methodology based on algebraic specification and a proof tool, the *HRP Prover*. One of the virtues of this methodology is that the same language, tool and proof techniques can be used both in specification and design, even down to a "concrete" specification which can be automatically translated into code. In the specification phase, the theorem prover is used to verify and validate the specification, while in the design phase the same tool is used to verify the correctness of the design steps.
- Testing a program means to execute it with selected test data to demonstrate that it performs its task correctly. Ideally the test data should be selected so that all potentially residual faults should be revealed. The Halden Project have performed several investigations of *testing methodologies*. An ongoing activity at HRP on testing is an experimental evaluation of a method, the so-called PIE (Propagation, Infection, Execution) method.

3.3 Future Work, Boundary Conditions and Needs

Maintenance and operating costs comprise a major portion of total cost and plant operators will be searching for means to enhance the plant availability through efficient control and surveillance systems. Currently, most plants operate with instrumentation and control systems for which industry support - and sometimes spare parts - are lacking or diminishing. The modernisation of existing control rooms will be a priority issue for the nuclear industry in the years to come. This will involve increased use of automation, compact operator workstations, databases, integration methods and digital display systems. Such programmes increase the need to develop guidelines and methods to facilitate the analogue-digital system transition.

Regulatory authorities are faced with the challenge of advanced information technology both in the control room and in the plant process and safety systems. The ability to analyse and anticipate the consequences of changes in operational practice or in the human-machine balance will remain an important focus of the human factors research. Compilation and analyses of international operational experience performed elsewhere can be a very valuable basis for understanding why errors are made, and for identifying both common elements and cultural and national differences. The need of proven methods for deriving validated conclusions from operational experience and HAMMLAB work should also be addressed. The experience from work with non-nuclear industries can be very beneficial for nuclear

control room applications, considering that other industries are often more advanced in the use of specific technologies.

3.4 Aims of the MMI Programme for the Time Period Between Year 2000 and 2002

Advanced computer-based human system interface technologies are being introduced into existing nuclear power plants to replace the existing interfaces. These developments can have significant implications on plant safety in that they will affect the overall function of the personnel in the system: the amount, type and presentation of information; the ways in which personnel interact with the system; and the requirements imposed upon personnel to understand and supervise an increasingly complex system.

The programme for 2000 - 2002 is intended to address the above issues by means of extensive experimental work in the human factors, control room design and computer-based support system areas. The work will be based on experiments carried out in the upgraded Halden Man-Machine Laboratory facility (HAMMLAB) which will become an even stronger nucleus of the research programme. The proposal is to a great extent based upon input from the Project's participants and contains four main areas of activity:

- *Experimental Programme and Operation of HAMMLAB.* The use of the laboratory will increase in terms of type and size of experiments, extended operational regimes and more realistic work settings. Installations of several advanced operator support systems on the new simulators are proposed to demonstrate the benefits of such systems in an integrated control room environment. An ambitious experimental programme is planned that will take extensive use for the *HAMMLAB* facility. The programme will address a wide range of issues including control room layout, interaction modes, information presentation and display design, levels of automation, human error and collaborative work.
- *Man-Machine Interaction* work aimed to extend the knowledge about the characteristics of human performance in process control environments and to demonstrate how this can be used in the specification and design of solutions to specific problems. The proposed programme addresses hybrid and advanced control rooms, contemporary and future man-machine interaction, interface design, individual and collaborative work, human error, function allocation and automation, and further method development.
- *Plant Performance Monitoring and Optimisation*, exploring and demonstrating system solutions that have potentials for improving plant performance and optimising plant operation as well as improving operational safety. The proposed activities comprise development of new and more robust support systems. Also, it is proposed to investigate how new technology as e.g. Virtual Reality can be used in operation and maintenance training.
- *System Safety and Reliability*, investigating the benefit of formal software development methods for computer systems with high reliability requirements. The integration of computer systems in plants makes it necessary to evaluate these in the total safety assessment context. It is proposed to study the incorporation of different evaluation methods for safety assessment of programmable plant control and supervision systems.

THE OECD HALDEN REACTOR PROJECT

FUELS TESTING PROGRAMME:

METHODS, SELECTED RESULTS AND PLANS

W. Wiesenack, T. Tverberg

Institutt for Energiteknikk
OECD Halden Reactor Project, Norway

26th Water Reactor Safety Information Meeting
Bethesda, Maryland, USA
26 - 28 October, 1998

ABSTRACT

The fuels testing programme conducted in the Halden reactor (HBWR) is aimed at providing data for a mechanistic understanding of phenomena which may affect fuel performance and safety parameters. It is based on more than thirty years of experience and the development of reliable in-core instrumentation, versatile irradiation rigs and loop systems for the simulation of light water reactor conditions.

The fuels performance studies focus on implications of high burnup. The instrumentation typically allows to assess thermal property changes as function of burnup, fission gas release as influenced by power level and operation mode, fuel swelling, and pellet-clad interaction. Relevant burnup levels (> 50 MWd/kgU) are provided through long term irradiation in the HBWR and through utilisation of re-instrumented fuel segments originating from commercial light water reactors. While UO₂ fuels still represent the majority of the test materials, other variants such as mixed oxide and Gd-bearing fuel receive increasing attention.

The thermal behaviour of uranium fuel as function of burnup has been investigated with a number of experiments which constitute a data base for the assessment of UO₂ conductivity degradation. The derived modification of UO₂ thermal conductivity is suitable for the explanation of temperatures measured in re-instrumented BWR fuel segments which have been further irradiated in the Halden reactor.

Various aspects of fission gas release are investigated with a number of experiments. The paper provides an example of release behaviour during normal operation as function of burnup and grain size. Regulations usually require that rod overpressure due to fission gas release does not lead to increased fuel temperatures due to clad lift-off and opening of the fuel-clad gap. The Halden Project is therefore conducting experiments to assess the cladding creep behaviour at different stress levels and to establish the overpressure below which the combination of fuel swelling and cladding creep does not cause increasing fuel temperatures.

Pellet-clad mechanical interaction (PCMI) is manifested with clad elongation measurements, and data originating from re-instrumented high burnup fuel are shown. The measurements provide information on the strain during a power increase, the relaxation behaviour, and the extent of a possible ratcheting effect during consecutive start-ups.

Further investigations as indicated above are planned in the current and next programme period from 2000 to 2002. It is foreseen to study the behaviour of mixed oxide fuel, Gd-bearing fuel and other variants developed in conjunction with burnup extension programmes. To this end, fuel segments irradiated in light water reactors and having reached high exposure have been procured. Some of them will undergo a burnup extension in the HBWR to reach burnups not yet achieved in LWRs, while others will be re-instrumented and tested for a shorter duration only. Plans are also being developed for testing high burnup fuel in power oscillation and LOCA conditions.

1. INTRODUCTION

Investigations of fuel performance in steady state and transient operation conditions have constituted a major part of the experimental work carried out in the Heavy Boiling Water Reactor (HBWR) at Halden since its start-up in 1959. The in-core studies were supported by the development and perfection of instrumentation and experimental rig and loop systems where reactor fuels and materials can be tested under PWR and BWR conditions^[1]. Fundamental knowledge and contributions to the understanding of LWR fuel behaviour in different situations could thus be provided in support of a safe and economic nuclear power generation.

Fuels testing at the Halden Reactor Project has for a number of years focused on implications of extended burnup operation schemes aimed at an improved fuel cycle economy. The experimental programmes are therefore set up to identify long term property changes with an impact on performance and safety. While PIE ascertains the state existing at the end of irradiation, in-core instrumentation provides a full description of performance history, cross correlation between performance parameters, on-line monitoring of the status of the test, and a direct comparison of different fuels and materials. Trends developing over several years, slow changes occurring on a scale of days or weeks, and transients from seconds to some hours can be monitored. The data generated in the fuels testing programmes originate from in-pile sensors which allow to assess:

- fuel centre temperature and thus thermal property changes as function of burnup;
- fission gas release as function of power, operational mode and burnup;
- fuel swelling as affected by solid and gaseous fission products;
- pellet - cladding interaction manifested by axial and diametral deformations.

The irradiation of instrumented fuel rods is carried out in specialised rigs according to test objectives, e.g. long term base irradiation, diameter measurements or ramps and overpower testing. In addition to fuel instrumentation, some rods in experimental rigs have gas lines attached to their end plugs. This allows the exchange of fuel rod fill gas during operation and makes it possible to determine gas transport properties as well as the gap thermal resistance and its influence on fuel temperatures. It is also possible to analyse swept out fission products for assessment of structural changes and fission gas release. This is an important experimental technique for the high burnup programmes currently being executed and defined for the period 2000 - 2002.

The examples of experimental work and results selected for this paper relate to high burnup fuel performance with respect to thermal behaviour, fission gas release, PCMI and cladding creep. They can be used for fuel behaviour model development and verification as well as in safety analyses.

2. SELECTED RESULTS FROM THE FUELS TESTING PROGRAMME

The examples discussed in the following sections represent only a fraction of the data base on fuel behaviour from zero to >90 MWd/kg burnup. While urania still represents the dominant fuel type, variants such as fuels with additives, Gd-bearing fuel and mixed oxide fuel receive increasing attention. It is therefore the aim to gradually build up a data base similar to the one existing for standard urania fuel.

2.1 Degradation of UO₂ thermal conductivity and thermal behaviour of high burnup fuel

A good knowledge of the fuel temperature is fundamental for fuel behaviour modelling since most properties and phenomena are temperature dependent. An accurate description of the temperature distribution in a fuel rod is therefore required before other effects can be quantitatively defined.

For high burnup fuel, several effects with an influence on thermal performance have been identified and made the subject of experimental work. One of the most important phenomena in this regard is the degradation of UO₂ thermal conductivity. Others are the changes induced by the formation of a porous rim and gap conductance as influenced by gap closure and fission gas release. These questions are addressed in the Halden Project experimental programme in separate effects as well as integral behaviour studies involving fuel with burnup from 50 to >90 MWd/kgUO₂.

Conductivity degradation has been manifested both with simulated^[2] and in-reactor burnup^[3] and is now generally accepted as an important phenomenon to be considered in modelling of high burnup fuel behaviour. The Halden Project's fuel testing programme contains a number of experiments where temperature measurements allow the conductivity degradation to be inferred. The evaluation of temperature changes with burnup has resulted in a modification of the MATPRO^[4] formulation for UO₂ conductivity:

$$\lambda = \frac{1}{0.1148 + 0.0035 \cdot B + 2.475 \cdot 10^{-4} \cdot (1 - 0.00333 \cdot B) \cdot T} + 0.0132 \cdot e^{0.00188 \cdot T}$$

with temperature T in °C, burnup B in MWd/kg/UO₂ and conductivity λ in W/mK for fuel of 95% t.d.

The modified formula is derived from in-pile data and therefore, in addition to the influence of fission products entrained in the fuel matrix, accounts for all other irradiation dependent effects which may have an influence on conductivity, i.e. microcracking, Frenkel defects and the formation of small fission gas bubbles.

Application to re-instrumented commercial fuel with high burnup

Fuel retrieved from LWRs or other types of reactors can be fitted with instrumentation, e.g. fuel centreline thermocouple, pressure transducer and cladding elongation detector. The re-instrumentation technique is well developed and has been applied to numerous fuel segments related to bilateral and HRP joint programme fuels testing. Relevant data from typical fuels with high burnup can thus be obtained without several years of waiting time normally required for burnup accumulation.

BWR fuel with a burnup of 59 MWd/kgUO₂ has been re-instrumented and then irradiated in the Halden reactor with the objective to study the thermal, fission gas release, and PCMI performance. PIE of sibling

fuel showed that a rim structure and a bonding layer had formed. The temperature data therefore reflect the combined influences of high burnup effects mentioned above: conductivity degradation, the thermal resistance of the porous rim with a maximum burnup of about 150 MWd/kgU, the gap between pellet and cladding, and eventually also fission gas release. The start-up temperature data are shown in Fig. 1 together with a code prediction assuming:

- conductivity degradation according to the model given above,
- periphery-peaked power and burnup distribution according to the TUBRNP model^[5],
- porosity distribution with a maximum at the periphery and decreasing to densified fabrication porosity for local burnup < 70 MWd/kgU (developed rim structure),
- gap closure at the power achieved at the end of BWR irradiation (12 kW/m).

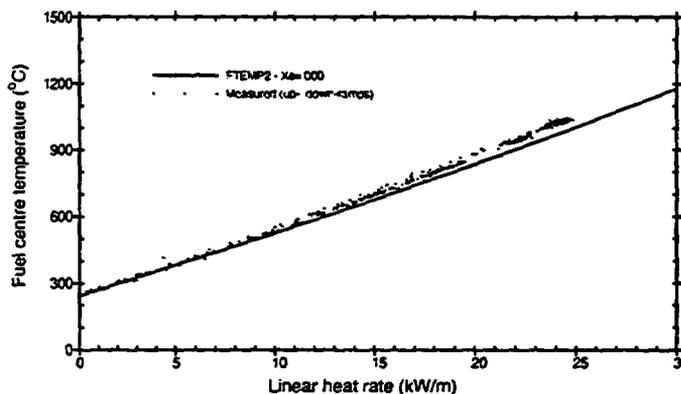


Fig.1 Temperatures measured in high burnup fuel and comparison with code prediction using the model for conductivity degradation

With these assumptions, a very satisfactory agreement between measured and calculated fuel temperatures can be obtained, confirming the validity of the conductivity degradation model for commercial fuel.

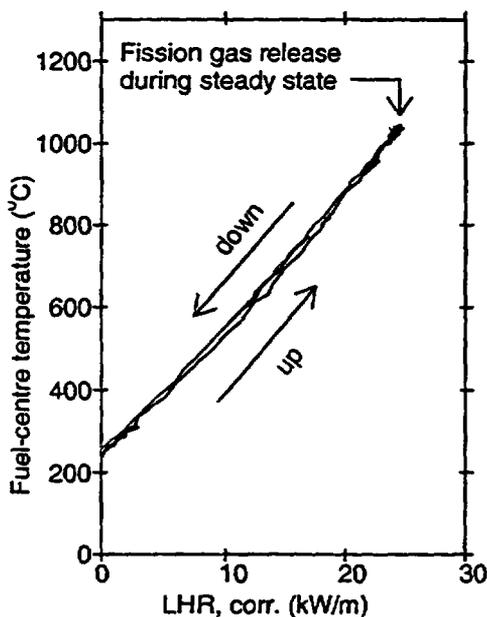


Fig. 2 Fuel temperature during a sequence of power increase, steady power with fission gas

Fission gas release during steady state operation eventually led to a 50% He / 50% FG mixture of gases in the fuel rod. However, this had little influence on gap conductance and fuel temperatures as is evident from the data shown in Fig. 2. The up-ramp (before FGR) and the down-ramp (after FGR) are virtually identical and it can be concluded that the gap conductance of fuel with high burnup is quite independent of the gas composition due to the tightly closed gap at power. This observation is confirmed by several other HBWR experiments where the fill gas can be exchanged in-pile through gas lines. An underestimation of gap conductance at high burnup will lead to overprediction of fuel temperatures (stored energy) and fission gas release and may thus severely impact safety assessments.

2.2 Fission gas release

The release of fission gas from UO_2 fuel continues to be a subject of considerable interest. At high burnup, the release may lead to rod overpressure and become a life-limiting factor. The influence on fuel temperatures and stored energy via gap conductance has direct consequences for the assessment of core reliability and safety during normal operation and transients.

Fission gas release from fuels with different grain size

In order to decrease fission gas release, fuels with large grains are being developed by vendors and tested in the Halden reactor. Also the joint programme addresses the mitigating effect of grain size on fission gas release with several experiments. For the one described below, fuel rods pre-irradiated in the HBWR were re-instrumented with pressure transducers with the initial objective to investigate the influence of power cycling operation on fission gas release. After establishing that power cycling did not cause enhanced fission gas release at the associated burnup level (30 - 40 MWd/kgUO₂), the irradiation continued normally to a burnup > 80 MWd/kgUO₂.

The two rods DH and DK differ with respect to grain size (6 and 17 μm) and gap size (200 and 360 μm). Calculating fission gas release from the change in rod pressure showed that the characteristics of release were different in the two rods. Whereas the release from rod DH with the small gap and small grain fuel increased rather gradually with time, that in rod DK showed a greater sensitivity to irradiation conditions. In particular, it showed a rapid release of fission gas during the initial rise to power.

The in-pile data have been analysed using a simple fission gas release model based on single gas atom diffusion with re-solution from grain boundaries. The predictions were benchmarked against the Halden empirical FGR threshold and compare very favourably with the data as can be seen from Figs. 3 and 4.

The conclusion reached from this experiment and code comparison is that the characteristics of the fission gas release observed do reflect the differences in rod design, namely, enhanced temperatures brought about by the large as fabricated gap and large grain size in rod DK compared with slightly lower temperatures and smaller grain size in rod DH. Of the two parameters studied in this experiment, the largest effect on FGR was induced by the difference in fuel-to-clad gap.

The further irradiation is of interest in light of plans to possibly achieve burnups of 100 MWd/kgU in commercial power reactors for even better fuel utilisation and waste reduction. The experiment has the potential to reach such a burnup level and to provide valuable lead data both from the in-core measurements and PIE.

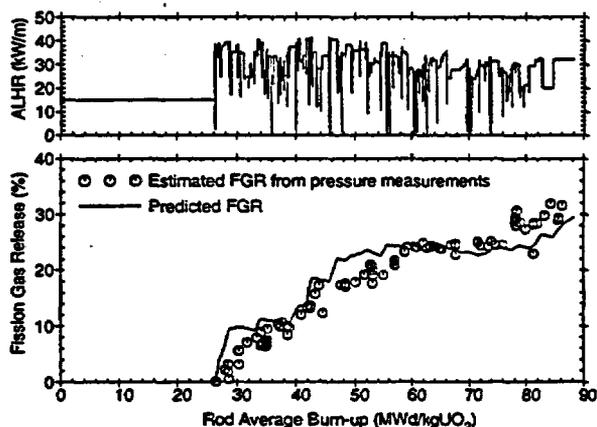


Fig. 3 Predicted and measured fission gas release (rod DH)

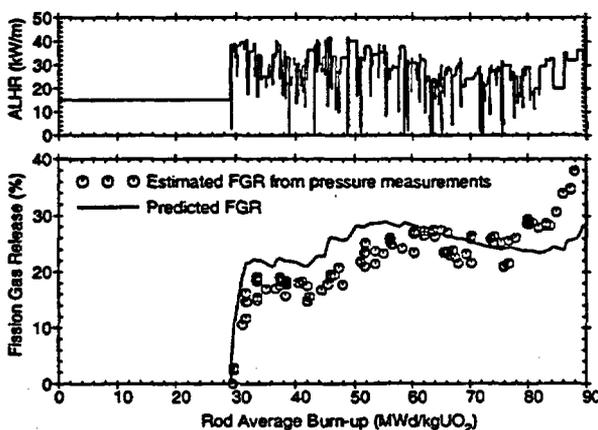


Fig. 4 Predicted and measured fission gas release (rod DK)

2.3 Pellet - clad mechanical interaction

Standard fuel designs employ a gap between pellet and cladding of about 2% of the pellet diameter. The gap closes gradually due to fuel swelling and cladding creep-down with the potential of increasing pellet-clad mechanical interaction (PCMI) with increasing exposure. This may pose restrictions on reactor operation with respect to rate and amount of power increases.

PCMI can be measured in-pile in two ways: with a diameter gauge moving along the length of a rod, and with a cladding elongation detector. The latter can also be fitted to pre-irradiated segments from commercial power reactor fuel rods. Cladding elongation data can be evaluated with respect to:

- onset of PCMI and amount of interaction;
- relaxation behaviour at constant power;
- permanent elongation due to overstraining (plastic flow), creep and growth;
- ratcheting interaction in conjunction with cyclic power changes.

The data of the following example are obtained from two 433 mm long re-instrumented PWR fuel rods with a final burnup of 37.5 MWd/kg UO₂ (42.5 MWd/kgU)^[6]. The rods, which differ with respect to fuel grain size (rod 1: 8.5 μm, rod 2: 22 μm), have also been used for fission gas release studies not reported here. Fig. 5 depicts the cladding elongation during periods of steady state operation (normalised to a constant power of 35 kW/m) for the entire time of irradiation. Several of the effects mentioned above can be identified:

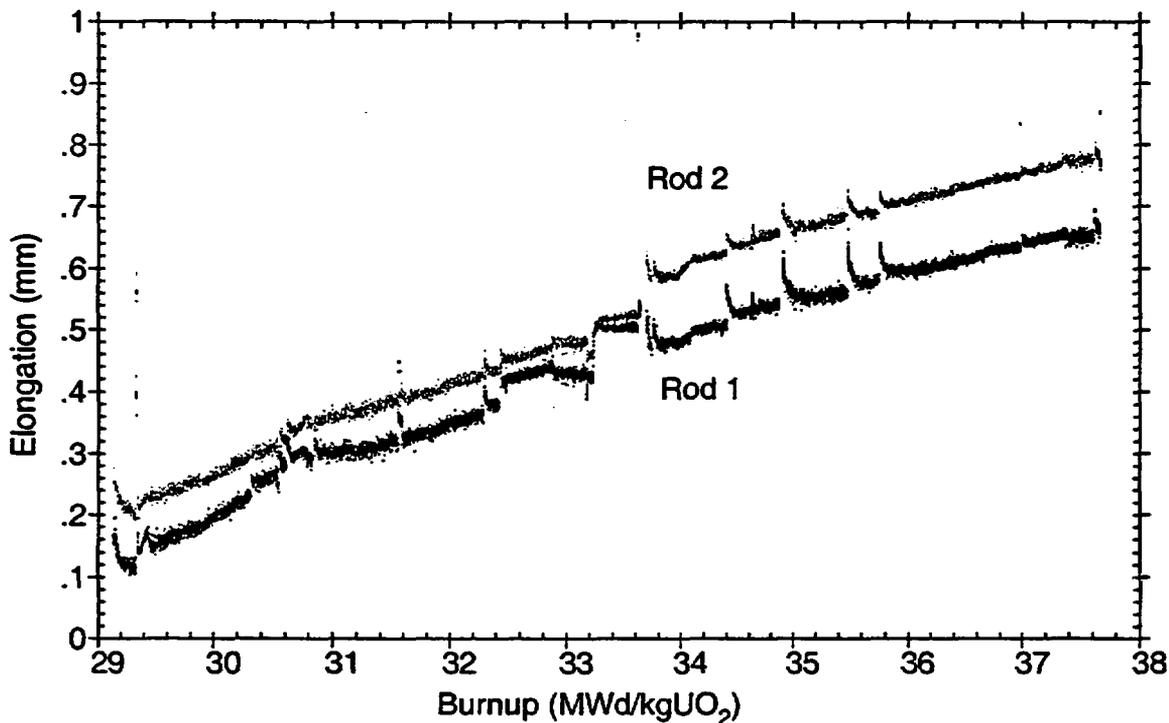


Fig. 5 Cladding elongation at steady state power of 35 kW/m for a PWR fuel rod with medium burnup

Ratcheting: The elongation peaks (especially visible for burnups > 33.5 MWd/kg UO₂) are associated with shut-down/start-up sequences. They indicate ratcheting, i.e. interaction onset during start-up occurs at somewhat lower power than the release of interaction (loss of pellet-clad contact) during the preceding power decrease. It appears that the small-grain fuel exhibits more ratcheting than the large-grain fuel.

Relaxation: The additional stress caused by ratcheting is relaxed by fuel creep within a few days. This is also the case for the first start-up.

Permanent elongation: The general trend of the elongation vs. burnup data indicates irradiation induced growth of the cladding. There are no obvious signs of plastic strain induced by the ratcheting elongation peaks.

From the example above, it should be clear that cladding elongation data contain a wealth of information. Gap closure, fuel-clad compliance and amount of PCMI are obvious subjects of analyses and data interpretations. A number of other effects and properties have been evaluated over the years by means of cladding elongation:

- thermal energy stored in a fuel rod;
- thermal conductivity of the zirconium oxide layer on the waterside surface of a fuel rod;
- swelling of high burnup fuel with bonding between fuel and cladding
- loss of contact between fuel and cladding due to rod overpressure and creep-out (noise analysis);
- relative difference of creep properties of UO₂ and MOX fuel.

Only few fuel modelling codes try to include pellet-clad interaction in a non-simplistic way, and the difficulties of modelling axial PCMI are recognised. But even without detailed model interpretation, cladding elongation data can and should be valued for providing insight into various aspects of fuel behaviour.

2.4 Cladding creep reversal and rod overpressure

Fission gas release at high burnup may result in the rod pressure exceeding the coolant pressure. A creep-out of the cladding may then open the fuel-cladding gap and lead to increasing fuel temperatures and further, increased fission gas release. In order to assess the consequences to fuel integrity, the creep characteristics of cladding material must be known.

Cladding creep data at high fluence in the presence of neutron flux were produced in the Halden reactor under representative LWR conditions in a diameter measurement rig. A gas line connected to the cladding tube enabled to change the rod inner pressure. In this way, several stress reversals were produced, and the cladding creep was measured in-pile by a diameter gauge with a relative precision of $\pm 2 \mu\text{m}$. Unlike PIE which only provides a single point, the results obtained show in a unique manner the development from primary to secondary creep.

The reaction of pre-irradiated BWR cladding material (fluence $6 \times 10^{21} \text{ n/cm}^2$, $E > 1 \text{ MeV}$) to stress reversals is shown in Fig. 6. The rod diameter changed in a stepwise manner whenever the applied stress was changed. Fig. 6 also shows that immediately following each stress change, although sometimes difficult to

detect, primary creep occurred. These data have provided insight in the reaction of cladding material to changing stress conditions and the relation of primary and secondary creep to stress change and stress level^[6]. They are used by Halden Project participants for modelling the creep-out behaviour at high burnup as consequence of rod overpressure. A question posed before conducting the test was whether primary creep would recur with every stress change. This was answered in a direct manner and the result has a bearing on the modelling of cladding failure induced by power changes.

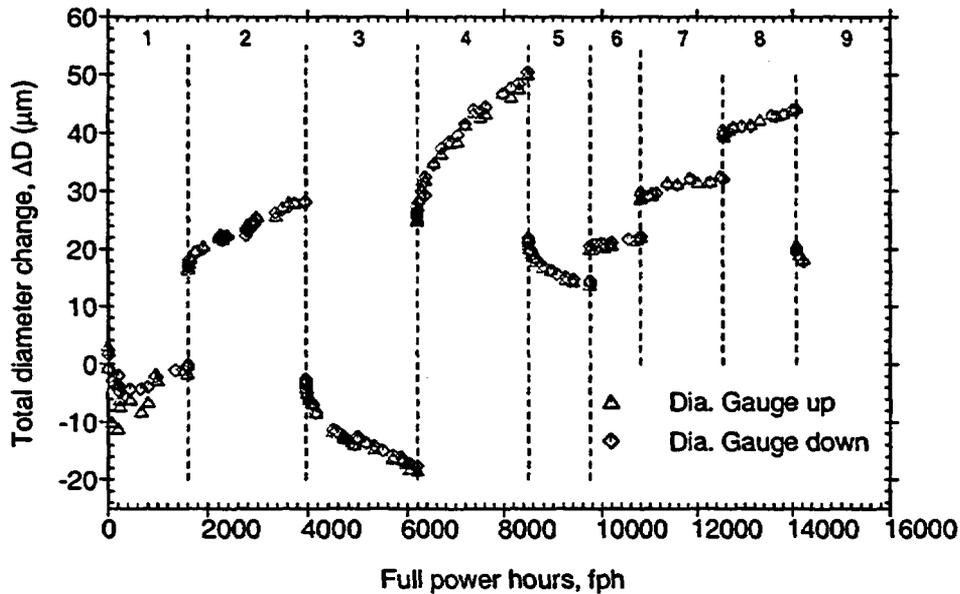


Fig. 6 Creep response of cladding tube subjected to compressive and tensile stress.
The recurrence of primary creep can be noted.

A complementary test, in which a pre-irradiated PWR rod equipped with a fuel centreline thermocouple is subjected to rod overpressure, was executed with the aim to determine the pressure beyond which the fuel temperature will increase due to clad creep-out. The fuel (burnup 55 MWd/kg) was re-instrumented with a fuel thermocouple and a cladding elongation detector. The rod overpressure was controlled with a high pressure gas supply system connected to the fuel rod with a gas line. This feature also allowed the exchange of fill gas ($\text{He} \leftrightarrow \text{Ar}$) during operation, hydraulic diameter measurements, and fission gas release analysis by means of gamma spectroscopy. The latter was also used to assess the contribution of fissions from the remaining U-235 and the Pu generated via conversion of U-238^[7]. The best agreement between the release-to-birth ratio of the Kr and Xe isotopes was obtained when the U/Pu composition as measured by PIE of sibling fuel was assumed for the fission yield. Unique temperature data in response to different levels of rod overpressure have been obtained, and it was found that overpressure >130 bar was required to produce increasing fuel temperatures as a result of clad lift-off and gap opening.

3. FUTURE ACTIVITIES RELATED TO FUEL PERFORMANCE INVESTIGATIONS

The demand for an economic and flexible, yet safe operation of nuclear power plants continues to pose considerable challenges. The continuation of the Halden Reactor Project experimental programme in the years 1999 - 2002 therefore focuses on the following issues related to fuel performance:

- Fuel high burnup capabilities in normal operating conditions, aiming at providing fuel property data needed for design and licensing in the burnup range 60 - 100 MWd/kg. Both test fuel and re-fabricated commercial fuels will be used, including Gd bearing and MOX fuels.
- Fuel high burnup capabilities in safety transients, aiming at providing experimental complements to investigations conducted elsewhere on loss of coolant and reactivity initiated transients. Tests on short-term dryout and possibly on power-coolant flow oscillations are also considered.

As before, the test execution will rely on experience with instrumentation, re-fabrication of irradiated fuel and cladding, experimental rigs providing suitable conditions, and the input from participants regarding issues and phenomena to be addressed as well as test implementation and execution details.

4. REFERENCES

- [1] *O. Aarrestad*: Fuel Rod Instrumentation; IAEA meeting on "In-core Instrumentation and In-situ Measurements in Connection with Fuel Behaviour", Petten, NL, 1992
- [2] *P. G. Lucuta, R. A. Verrall, H. Matzke, I. J. Hastings*: Thermal conductivity and gas release from SIMFUEL; IAEA Technical Committee Meeting on fission gas release and fuel-rod chemistry related to extended burnup (IAEA-TECDOC-697), Pembroke, Canada, 1992
- [3] *E. Kolstad, C. Vitanza*: Fuel rod and core materials investigations related to LWR extended burnup operation; Journal of Nuclear Materials 188 (1992), pp 104-112
- [4] *D. L. Hargman, G. A. Reymann (eds.)*: MATPRO version 11-A Handbook of Materials Properties for Use in the Analysis of Light Water Reactor Fuel Rod Behaviour. TREE-NUREG-1280, February 1979
- [5] *K. Lassmann, C. O'Carroll, J. van de Laar, C. T. Walker*: The radial distribution of plutonium in high burnup UO₂; Journal of Nuclear Materials 208 (1994), pp 223-231
- [6] *T. Tverberg*: Studies of PCMI from cladding elongation measurements performed in the HBWR; IAEA Technical Committee Meeting on "High burnup fuel specially oriented to fuel chemistry and pellet clad interaction", Nyköping, Sweden, September 1998
- [7] *M. A. McGrath*: In-reactor creep behaviour of Zircaloy-2 under variable loading conditions; Meeting of the German Nuclear Society, Deutsches Atomforum, Munich, May 1998
- [8] *R. J. White*: The measurement of radioactive fission gas release. HRP internal note, May 1998.

ACHIEVEMENTS AND FURTHER PLANS FOR THE OECD HALDEN REACTOR PROJECT MATERIALS PROGRAMME

T.M. Karlsen

Institutt for energiteknikk

OECD Halden Reactor Project

P.O. Box 173, N-1751 Halden, Norway

Tel. + 47 69212200; Fax: + 47 69212201

ABSTRACT

The materials programme at Halden, in addition to cladding corrosion studies, is aimed also at addressing the effects of operating conditions and water chemistry variables on core materials behaviour, particularly as related to reactor pressure vessel integrity and Irradiation Assisted Stress Corrosion Cracking (IASCC), the materials degradation phenomenon which affects the structural integrity of stainless steel and nickel based components. The aim of the experimental work is to improve the understanding of materials ageing processes, to demonstrate the benefits of mitigation measures and to evaluate properties of materials which have been subjected to long in-reactor service. While a number of the studies are performed in loops which simulate light water reactor environments in terms of thermal-hydraulic, radiation and water chemistry conditions, dry irradiation facilities are also utilised, particularly in relation to studies aimed at determining the effects of fluence on material integrity.

1. INTRODUCTION

The objectives of the activities in the materials programme are to improve the understanding of materials ageing and degradation processes as well as to demonstrate methods designed to increase component lifetime, all of which are of considerable importance as safety related issues. The focus is on in-reactor experiments addressing core and vessel materials and, to this end, a range of different studies are being conducted. Under the Irradiation Assisted Stress Corrosion Cracking (IASCC) test programme, the IASCC susceptibility of stainless steels (and nickel base alloys), the effects of radiation and water chemistry on IASCC, the quantification of crack growth rates as a function of stress intensity and the benefits of mitigation measures such as HWC have been addressed. Under dry irradiation conditions, the effects of fluence on the microstructural and mechanical properties of ferritic and austenitic reactor internals materials are being studied.

Aspects of each of these areas of research are described in further detail in the sections which follow.

2 IRRADIATION ASSISTED STRESS CORROSION CRACKING (IASCC) PROGRAMME

2.1 Crack Growth Studies

In the series of crack growth studies that have been performed at Halden to date, several objectives have been addressed. The purpose of the early studies was to develop specimens and instrumentation methods that could be employed in monitoring the cracking behaviour of core component materials during exposure in representative LWR (typically BWR) environments. By utilising instrumented specimens, the impact of, for example, changes in water chemistry and / or applied stress intensity levels may be monitored on-line during the course of irradiation, as opposed to relying on post irradiation examination to determine behaviour.

In the first investigations, the possibility of using wedge-loaded Double Cantilever Beam (DCB) specimens, instrumented for crack propagation monitoring with the DC potential drop technique, as in-pile crack growth sensors was successfully demonstrated and the benefits of hydrogen water chemistry (HWC) in controlling cracking in both thermally and radiation sensitised stainless steels were confirmed. In a follow-on qualification study, instrumented, actively loaded DCB and miniaturised Compact Tension (CT) specimens were developed for crack growth versus stress intensity experiments. The specimens were equipped with pressurised bellows which enabled on-line control and variation of applied load. Typical examples of the crack growth rates recorded in the specimens as a function of stress intensity varying bellows pressure (stress intensity level) are presented in Fig. 1.

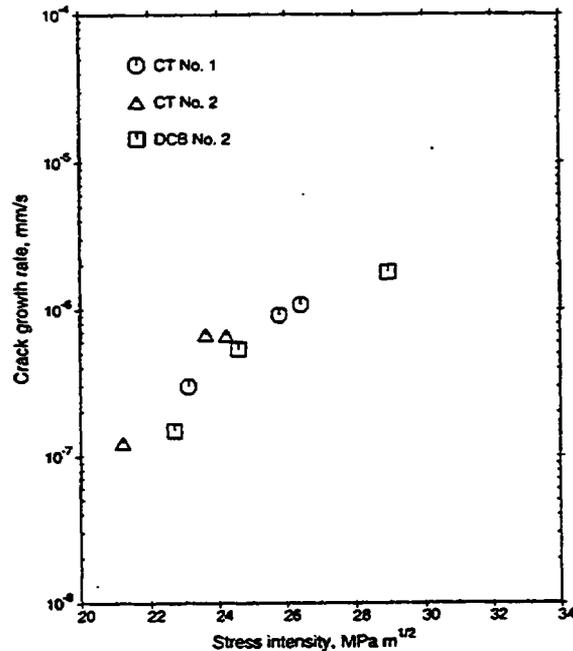


Fig. 1 Crack growth rates measured on CT and DCB specimens as a function of stress intensity introduced with pressurised bellows.

Following successful qualification of the instrumentation and loading techniques required for the crack growth studies, the second phase of the programme has addressed the feasibility of incorporating small sections of irradiated material into the crack growth region of the miniaturised CTs. The reconstitution technique has opened up the possibility of studying the crack growth behaviour of materials retrieved from irradiated structural components. Minimising the size of the test piece has advantages both in terms of facilitating specimen preparation (reduced activity levels) and also in overcoming limitations related to material availability (i.e. the amount of irradiated material required to produce the reconstituted CTs is minimal compared to the quantities that would be required to produce full-size CTs (or DCBs)).

The main objectives of the phase two studies, which utilise bellows-loaded, instrumented CTs with irradiated inserts, are to produce reliable crack growth rate data for irradiated core component materials.

The first of these investigations, which doubled as both a crack growth study and a feasibility study aimed at demonstrating the satisfactory performance of the reconstituted CTs, was completed in May 1998 after two irradiation cycles (~200 full power days) in the Halden reactor.

The rig contained 4 specimens, each loaded with bellows and instrumented for crack growth monitoring with the DC potential drop method. Irradiated material, in the form of a disc or a square (Fig. 2), was electron beam (EB) welded into the crack growth region of each CT. Discs were prepared from a 304 SS control blade handle (fluence $9 \times 10^{21} \text{ n/cm}^2$), a 347 SS tensile specimen (fluence $1.1 \times 10^{21} \text{ n/cm}^2$) and from a solution annealed 304 SS DCB (fluence $0.82 \times 10^{21} \text{ n/cm}^2$) which had been used in one of the earlier IASCC studies at Halden. An irradiated "square" was also prepared from the solution annealed 304 SS DCB.

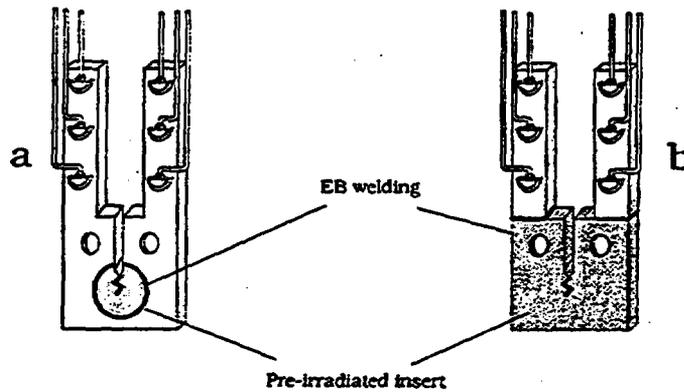


Fig.2 Miniaturised CTs with inserts of irradiated material with (a) disc and (b) square geometry

During irradiation, the specimens were loaded to stress intensities ranging from 16 to 22 $\text{MPa}\sqrt{\text{m}}$ and crack growth rates were calculated for each of the CT specimens during operation with NWC (~400 ppb O_2) and with HWC (~400 ppb H_2). The cracking rates recorded for each of the four specimens in NWC and HWC are summarised in Fig. 3. The ranking in crack growth rates appears reasonable, with the highest fluence material exhibiting the highest rate of crack propagation. Subsequent testing, also in alternating NWC and HWC conditions, indicated similar trends in behaviour although hydrogen additions appeared more effective in suppressing cracking in the lower dose than in the high dose material, a result that is to be studied further in new investigations.

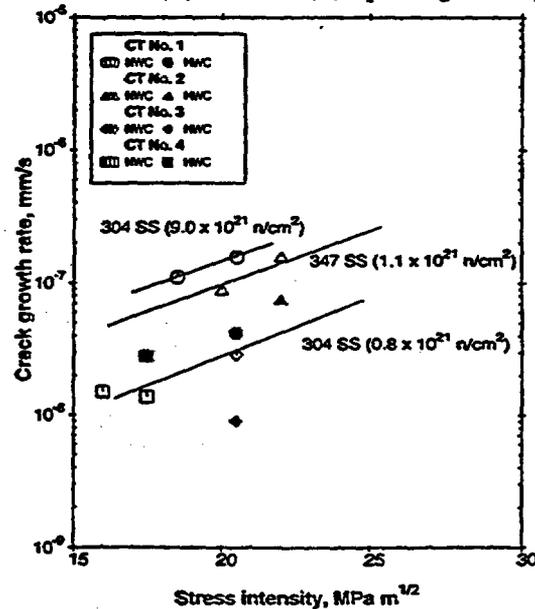


Fig. 3 Crack growth rates measured in irradiated CTs at different stress intensities and in NWC and HWC.

The test rig for the second of the phase two crack growth studies, which will be loaded at the beginning of 1999, also contains 4 reconstituted CT specimens (Fig. 4), each equipped with bellows and instrumented for crack growth monitoring with the DC potential drop method. Two of the test specimens are prepared from 347 SS top guide material (fluence of $\sim 1.5 \times 10^{21}$ n/cm²) from the Wurgassen nuclear power plant. One of the specimens is prepared from the high ($\sim 9 \times 10^{21}$ n/cm²) 304 SS control blade handle material and the fourth specimen, with a fluence of 0.9×10^{21} n/cm² is prepared from 316NG.

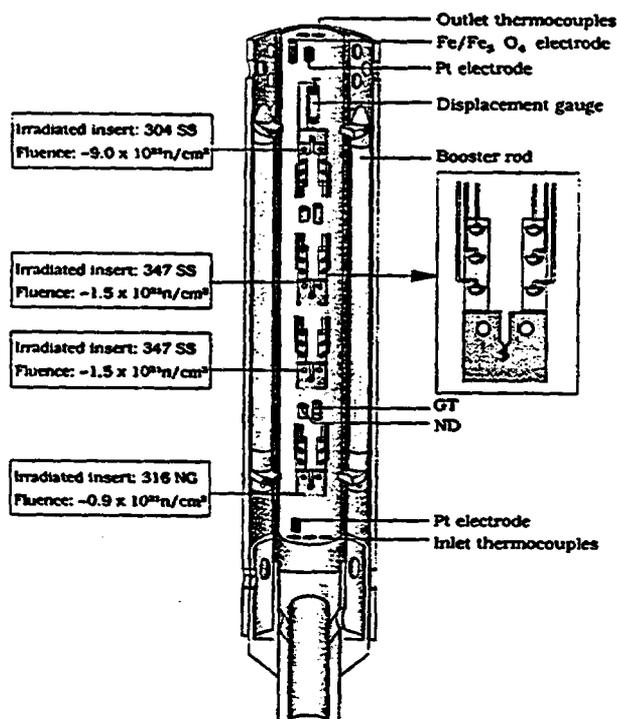


Fig. 4 Illustration of the reconstituted CT arrangement in the new crack growth rate study which commences in early 1999.

The matrix outlined above will:

- provide additional crack growth rate data on the high fluence control blade handle material
- enable comparison of the cracking behaviour of different materials (316 NG vs 347 SS) with similar fluence levels
- allow reproducibility in cracking behaviour to be determined by comparing the crack growth rates exhibited in two specimens prepared from 347 SS.

As in the preceding study, stress intensity levels in the range 15-20 MPa√m will be employed in the investigation and crack growth rates in the specimens will be measured under both NWC and HWC conditions.

Future Plans

The future experimental programme will continue to address the effects of typical BWR conditions on cracking behaviour in reconstituted CTs, with the studies aimed specifically at generating crack growth rate data over long time intervals, focusing, in particular, on materials with high fluence levels ($> 6 \times 10^{21}$ n/cm²). Future studies will also be aimed at determining if IASCC countermeasures, such as the introduction of HWC in BWRs, still remain effective for high dose materials representative of those

found in older plants. The benefits of noble metal or other coatings in enhancing the effects of hydrogen additions are also to be addressed. Finally, an experimental series devoted to measuring crack growth rates in core structural materials representative of those found in PWRs, where higher end-of-life fluences are reached has been initiated.

2.2 Crack Initiation Studies

In a crack initiation study, the effects of fluence and stress on the initiation of cracks in pressurised tube specimens is being evaluated. It is anticipated that results from this investigation will provide information on the boundary conditions which may result in the initiation, as opposed to growth, of cracks in structural component materials. The specimen geometry comprises tubes that are pressurised with argon gas to stress levels ranging from 0.8 to 2.75 times the yield stress of the unirradiated material. Of the 36 tube specimens in the test matrix, a total of 26 are prepared from sensitised 304 SS and the remaining 10 are prepared from cold worked 347 stainless steel and solution annealed 304 and 316L. The specimens are arranged in 6 strings, each with 6 tubes. Three of the strings are located in a fast neutron flux region in the test facility and the remainder is installed in a low flux position (Fig. 5).

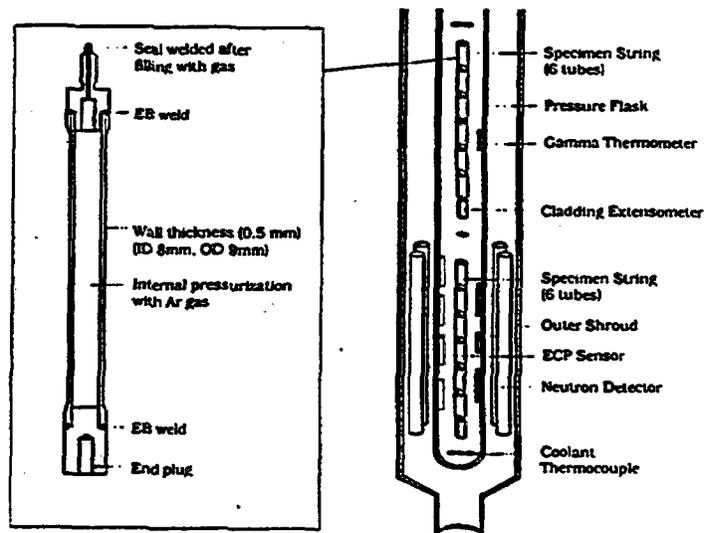


Fig.5 Arrangement of pressurised tube specimens in crack initiation study

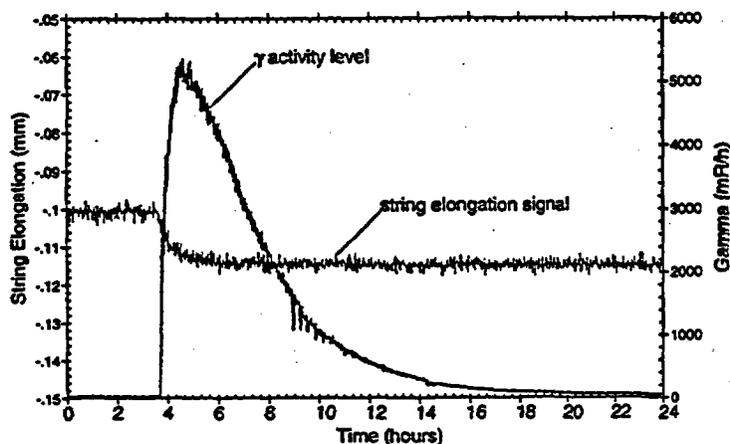


Fig. 6. Example of pressurised tube rupture detected on-line with string elongation and gamma activity measurements.

Tube ruptures are determined on-line by means of string elongation detectors and by increased activity levels which are measured in the coolant as a result of the release of the tube filler gas (activated Ar-41) into the water. A typical example of the signal changes occurring in the event of a tube rupture is illustrated in Fig. 6. Fluences in excess of 1.6×10^{21} n/cm² have been accumulated by the high flux specimens and irradiation is being continued to higher dose levels.

Future Plans

In addition to the pressurised tube geometry for crack initiation studies, alternative crack initiation specimens are being developed in collaboration with participants. The geometry, coupled with self-contained loading units and in-situ failure detectors, would make the arrangement suitable for installation also in commercial plants as a surveillance unit, since no pressure boundary penetrations are required.

3. DRY IRRADIATION PROGRAMMES

Together with the test programmes dedicated to studying the crack growth and crack initiation behaviour of core component materials, a series of dry irradiation programmes are also conducted in the Halden reactor. In this area, the test materials are represented by both core structural and reactor pressure vessel (RPV) materials. The test specimens, which take the form of, for example, TEM foils, Charpy, Tensile and / or CT specimens, are enclosed in capsules designed so as to meet specified temperature requirements and suitable reactor positions as selected so as to achieve the desired thermal and fast neutron flux conditions for the individual studies. Thermocouples inserted in the irradiation capsules may be used to record specimen temperatures and flux wire monitors are used in determining thermal and fast neutron fluence.

3.1 Irradiation of Reactor Pressure Vessel Materials

Neutron embrittlement effects on pressure vessel materials remain a safety issue which, traditionally is assessed on the basis of testing Charpy-type specimens irradiated as part of surveillance programmes. While limited material availability is being resolved through reconstitution techniques and miniaturisation of test specimens, further experimental verification is required in order to establish correlations between data from specimens with subsize and standard geometry. This is of importance in relation to the establishment of fracture toughness data and in assessments of the effects of pressure vessel annealing and subsequent re-embrittlement.

In a study which is scheduled to commence shortly, a number of standard Charpy V notch specimens, prepared from a typical WWER-440 weld material, are to be irradiated in dry conditions in the Halden reactor. The test, which is being conducted in co-operation with a participant, is aimed at complementing an IAEA round robin exercise on pressure vessel ageing. After the first irradiation phase, a number of the specimens are to be shipped to a specialised laboratory for mechanical testing, and the remainder will be annealed and re-irradiated before finally being discharged. The main objectives of the programme will be to study the effects of irradiation embrittlement, post irradiation annealing recovery and post-annealing re-embrittlement on material toughness.

Future Plans

For future experimental work in this area, it is anticipated that co-operation in international programmes in support of correlating data measured with Charpy specimens of varying size will continue. It is envisaged that the Project may contribute by providing irradiation at representative temperature and flux conditions in the Halden reactor, while collaborative arrangements for mechanical testing are established with participating organisations.

3.2 Irradiation of Model Austenitic Stainless Steels

In the case of dry irradiation of austenitic and Inconel structural materials, the objectives are to systematically evaluate the effects of variants such as chemical composition, heat or mechanical treatment on the irradiated material characteristics. The results from these studies, too, are of considerable importance in gaining a better understanding of the key factors affecting degradation of reactor components. In one such investigation, conducted in collaboration with Argonne National Laboratory and sponsored by the USNRC, the objective is to study, in a systematic manner, the effect of fluence on the microstructural and mechanical properties of austenitic stainless steels. In the investigation, a total of 96 Slow Strain Rate Tensile (SSRT) and 24 Compact Tension Fracture Toughness (CTFT) specimens, prepared from 27 model austenitic stainless steels (22 Type 304, 3 Type 316 and 2 Type 348) are being irradiated (under dry conditions) to three different fluence levels in the Halden reactor.

Discharged specimens are shipped to Argonne National Laboratory where post irradiation characterisation is undertaken. To date, half of the specimens (with low (0.4×10^{21} n/cm²) and medium $\sim 1.0 \times 10^{21}$ n/cm²) fluence accumulation) have been discharged. SSRT tests in high temperature BWR water (with 8 ppm oxygen) have been performed on the low fluence specimens and the fracture surfaces have been characterised by SEM (1,2). Testing of the medium fluence samples is in progress and the remaining 48 SSRT and 12 CTFT specimens are being irradiated to a higher fluence (2.5×10^{21} n/cm²). The final results, from all the tests, will be used to establish a database on susceptibility to IASCC, crack growth rate and fracture toughness of the irradiated materials. The data will also be used in conjunction with results of investigations on service-degraded LWR core internals in order to identify mechanisms of IASCC and to establish a method for evaluating the long-term structural integrity of core internal components.

Future Plans

A phase two dry irradiation programme is planned, aimed at further elucidation of the more recent findings obtained through examination of core component materials retrieved from operating reactors. The test materials to be studied in the second programme include alloy types 304, 304L, 316 and A690, again in the form of SSRT and CTFT specimens. The main objectives of the phase two programme will be to

- continue the process of identifying key impurities and fabrication factors influencing IASCC
- confirm the IASCC resistance of selected heats of material
- evaluate core-internal weld performance
- obtain a crack growth rate data base for austenitic steels and Alloy 690

Loading is scheduled for 1999 and the specimens will be irradiated under dry conditions to a fluence of 1.2×10^{21} n/cm².

REFERENCES

- (1) Environmentally Assisted Cracking in Light Water Reactors, Semiannual Report, July 1996-December 1996, NUREG/CR-4667, Vol. 23, ANL-97/10 pp 32-35
- (2) Environmentally Assisted Cracking in Light Water Reactors, Semiannual Report, January 1997-June 1997, NUREG/CR-4667, Vol. 24, ANL-98/6 pp 34-39

**ACHIEVEMENTS AND FURTHER PLANS
FOR THE OECD HALDEN REACTOR PROJECT
MAN-MACHINE SYSTEMS PROGRAMME**

Fridtjov Øwre

OECD Halden Reactor Project

P. O. Box 173, N-1751 Halden, Norway

Tel: + 47 69212200; Fax: + 47 69212201

ABSTRACT

The OECD Halden Reactor Project is a joint undertaking of nuclear organisations in 20 countries sponsoring a jointly financed research programme under the auspices of the OECD - Nuclear Energy Agency. The organisations participating to the Project represents a complete cross-section of the nuclear community, including regulatory bodies, vendors, utilities and R&D organisations. The programme is renewed every third year. The three main research areas at the Halden Project are : Fuels, Materials and Man-Machine Systems (MMS).

The research and development efforts in the MMS area were initiated on the basis of the experience gained through the Halden reactor dynamics experiments and the use of in-core instrumentation. Initially, efforts were spent on practical demonstrations of advanced concepts for closed loop control and core power distributions. Plant load-follow control were later successfully demonstrated on the Halden reactor. The Halden Project has since 1970, through international co-operation, successfully conducted research and development in areas related to control room systems, technology development as well as human factors. Since 1983 the work has utilised the experimental control room HAMMLAB linked to a full scope PWR simulator. A process is soon completed to provide this facility with new full scope simulators for BWR, PWR and VVER reactors, complemented with a Virtual Reality laboratory.

The programme has in later periods addressed the research needs of the nuclear industry in connection with introduction of digital I & C systems in NPPs and it has provided information supporting design and licensing of upgraded, computer-based control room systems, and demonstrated the benefits of such systems through test and evaluation experiments in HAMMLAB and in pilot installations in NPPs.

This paper describes *the facilities* used for the MMS research at the Halden Project followed by an overview of some recent achievements in the four main areas of research: *human factors research, experimental control rooms, plant surveillance and operations systems and enhancement and assessment of system quality.*

FACILITIES

HAMMLAB

The Halden Man-Machine Laboratory, HAMMLAB, was established in 1983 in order to serve as the main environment for performing realistic experiments within the MMS research area. Since its establishment, HAMMLAB has been the experimental focal point of the research within Human Factors, as well as the main test bed for computerised operator support systems being developed both at the Halden Project and at members organisations.

The NORS full-scope simulator has since the establishment of HAMMLAB been the laboratory's simulator basis. NORS is based on the Loviisa nuclear power plant in Finland.

The facility has three major functions:

- *process operation* - in the control room - is where operators (test subjects) are monitoring and controlling the NORS process in normal and disturbed plant conditions,
- *experimentation* - in the experimenters gallery - is where the experimenters set up, monitors and control the experiments, and where a database is collected during experiments consisting of human performance measurements as well data from the control room operation and the process itself,
- *evaluation* - in the experimenters gallery - where the experimenters analyse results from the experiments by means of various techniques and statistical packages.

HAMMLAB has undergone major upgrades and improvements since 1983, the last major one being performed in 1996 (1) with the introduction of a new unified human-machine interface and a new control room set-up. The upgrades have partly been made to support the Halden Project's research programmes, partly due to specific requirements set forth in bilateral funded experiments and studies.

HAMMLAB AS OF 1998

The studies being performed in HAMMLAB are of different size and complexity, ranging from large scale human factors experiments to small scale studies and tests. Due to this, the requirements to the laboratory vary a lot, and a flexible infrastructure is a necessity. The current control room is equipped with two operator stations and one supervisor station, as indicated in Fig. 1. All stations are situated on desks having wheels in order to ease shuffling around and varying the degree of compactness of the control room. In this way it is rather easy to restructure the control room for one or more operators. Another key issue is that all information is available on all screens, thus allowing for tests using single operators and few screens, or full shift crew.

The latest device introduced in HAMMLAB is a large interactive overview display, see also Fig. 1. Such devices are now being introduced in many control rooms for complex processes and it is a defined need to investigate what the content should be on such overview displays and what makes it a common reference point for the whole operating crew.

The carrying through of a large human factors experiment requires careful preparation prior to the actual execution of the experiment, and a large period for data analysis after the experimental execution. The data collection phase, i.e. the actual execution of the experiment, requires the availability of advanced data recording equipment. Audio and video recorders, eye movement tracking devices, computerised data logs of various kinds, are heavily used in addition to questionnaires and on-line expert commenting.

HAMMLAB of today provides the experimenters with advanced data recording equipment and everything is configured and operated from a specially designed experimenters' gallery. Fig. 2 shows a picture of the experimenters' gallery where the human factors experts carefully follow what is going on in the HAMMLAB control room.

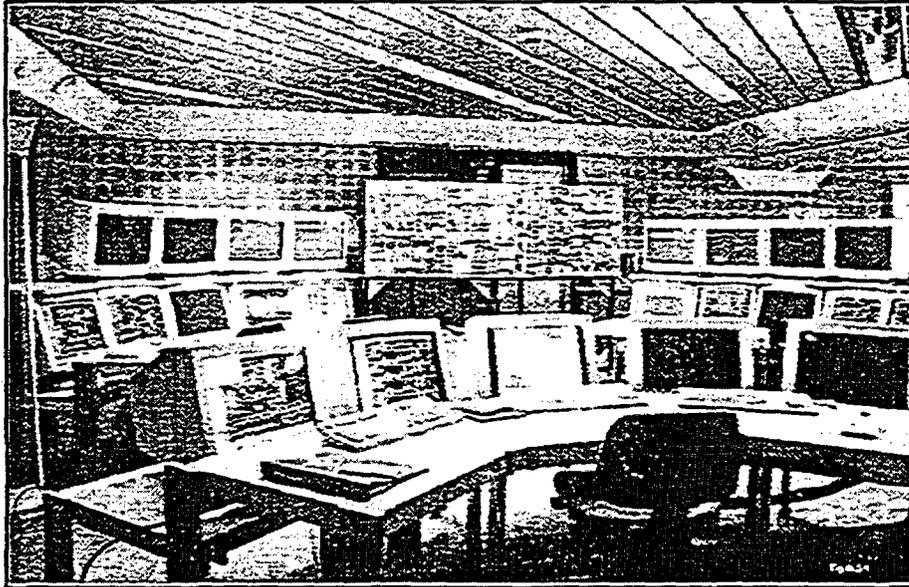


Fig. 1 The 1998 HAMMLAB Control Room

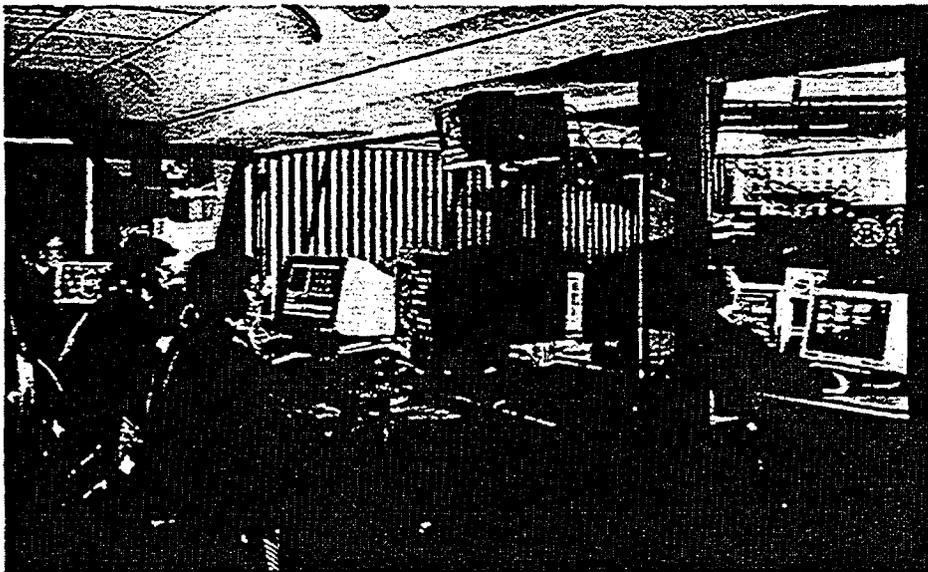


Fig. 2 The Experimenter's Gallery of HAMMLAB

HALDEN VIRTUAL REALITY CENTRE (HVRC)

Virtual Reality (VR) is an exciting new technology with many possible applications. VR has been defined by (2) as "A computer system used to create an artificial world in which the user has the impression of being in that world and with the ability to navigate through the world and manipulate objects in the world". Users of VR today reside in simulation, telerobotics, medicine, architecture, and entertainment and many other areas.

HVRC is a new complementary extension to HAMMLAB. The activities in the VR-centre are connected to on-going research at the Halden Project, but mostly to development projects for the industry such as nuclear, oil production, maritime, air-traffic control and process control. The VR-centre expands the possibilities to develop and evaluate new systems and methods for HMI. VR is being used to guide designers and inspectors in cost effective control centre design development and evaluation (Human Factors V&V); and to develop computerised training tools for outage maintenance and operational tasks.

Virtual Mock-Up's: VR as an engineering tool when modernising control rooms

As to-days control rooms needs upgrading from technological reasons there is now an opportunity to initiate the redesign by applying VR-technology. Our product is a "Virtual mock-up" where the end-users (the operators), engineers and managers with assistance of human factors specialists - together can formulate the optimal solution for their new control room. The possibility to actually "walk inside" the planned control room - and consider, modify, try out, discuss, agree on the location of desks, keyboards, CRTs and control boards with their instruments - from every angle and position before any design decision has been made - is a tremendous advantage. The final output from this exercise is a set of CAD drawings to the vendor actually building the control room.

The Halden Project is using this approach in several redesign projects for Swedish nuclear control rooms (3).

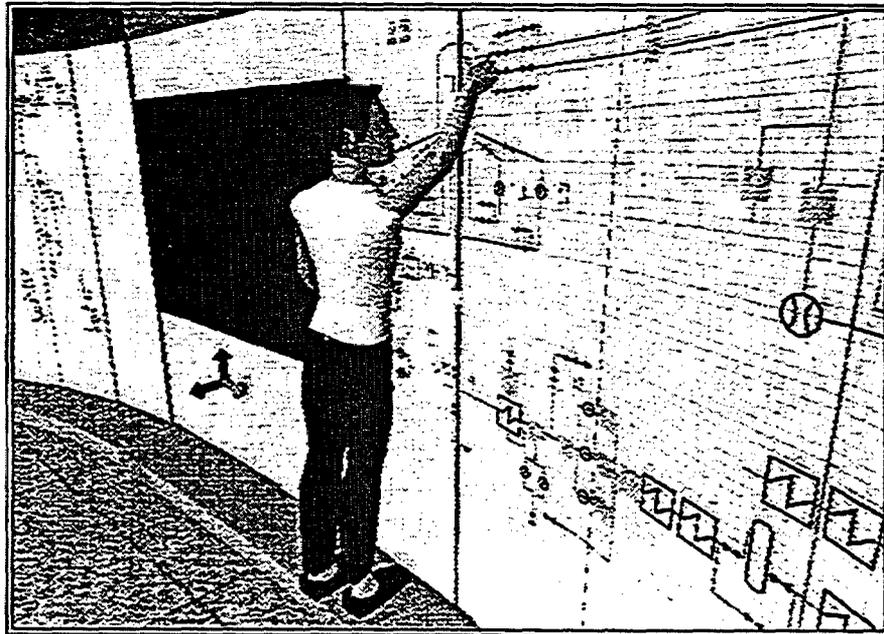


Figure 3. VR used in control room design

HUMAN FACTORS RESEARCH

The work on *human factors research* aims at providing knowledge about the capabilities and limitations of the human operator in a computerised control room environment. Understanding the impact of new technology on the role and performance of operating personnel is crucial in decision making concerning safety of nuclear power plants.

Experiments are carried out to assess operator performance in varying plant conditions and in presence of various support systems. Important issues which are addressed include mechanisms affecting human error and methods for its prediction; effects of alarm processing and presentation; effects of level of automation on operator performance as well as studies of operator performance at night.

At the Halden Project this understanding is accomplished through a combination of activities addressing operators cognition and information processing methods in various control room work situations, function and task allocation methods and test and evaluation of advanced support systems. In addition to providing knowledge about human performance, the program also gives results on the new methods and measures for studying human performance.

The program is divided in three activities: i) operator cognition and information processing, ii) test and evaluation and iii) methodological development

STUDIES OF OPERATOR COGNITION AND INFORMATION PROCESSING

Human Error

The aim of the activities on human error analysis project (HEAP) is to provide improved understanding of how operators diagnose disturbances and to identify potential errors or inefficiencies which may occur in the diagnostic process. The long term goals are to produce practical knowledge for system and man-machine interface design and to achieve better modelling of cognitive errors for representation in probabilistic safety analysis.

A series of pilot studies and one major experiment have been performed to find a reliable and valid methodology to investigate human error in a control room setting. Practical insights from these studies can be found in (4) and relates to i) diagnostic strategies and styles that have been observed in single operator and team based studies; ii) the qualitative aspects of the key operators support systems, namely MMI interface, alarms, training and procedures, that have affected the outcome of diagnosis; and iii) the overall success rates of diagnosis and the error types that have been observed in the various studies.

Complexity Profiling

The human error project has for some time been investigating what makes certain fault scenarios difficult for operators. One line of research has been to develop a questionnaire to measure diagnostic complexity.

Another line was a literature review and analyses of two questionnaire studies. The literature review was chosen to find out the factors or components thought to contribute to difficulty in diagnosis and problem-solving. Two experimental studies of complexity were carried out using two versions of a questionnaire based on the review. The studies were simulator based, using scenarios designed to be challenging and using PWR reactor staff.

The main goal was to identify dimensions of complexity and to study the relation between complexity and operator behaviour. The preliminary results indicate that further developments of the measures have potential for enhancing the understanding of complexity and how it is related to operator performance.

Error Recovery

One of the issues addressed by the current set of experiments is the question of detection and recovery of erroneous actions in the situation where they actually take place. Since detection clearly is a prerequisite for recovery, the experiments will in the first instance consider how well people are able to detect the erroneous actions they make and how the level of detection depends on the circumstances or conditions, such as interface, team, workload, etc. This also suggests that a possible objective for the next experiment is to study the conditions for recovery, as well as in more detail the specific circumstances that can facilitate detection.

Error Mode Prediction (CREAM - Cognitive Reliability and Error Analysis Method)

The main purpose of the first main HEAP experiment (autumn 1996) was to develop a method for predicting performance failures, specifically the error modes that can be expected for a specific task.

CREAM (Cognitive Reliability and Error Analysis Method) takes as input a detailed task description of a set of scenarios for an experiment. This description is analysed to characterise the nature of the individual task steps, as well as the conditions under which they must be carried out. Each task step is examined using a systematic classification scheme of possible error modes, and the likely error modes are identified. This gives a qualitative prediction of the possibilities for erroneous actions in a given task.

The 1997 Experiment further refined this method of prediction, which is a variation of the basic method of CREAM. OPAS (see below) was used as a main source for both the predictions of error modes and for analyses of the observed errors during the experiment.

Human Centred Automation

Professor Sheridan has in (5) said: "It has become evident that humans, when put in the role of monitor, supervisor and automation back-up in case of failure, may not perform. Humans become bored and un-alert during the long period when the automation does work. They may lose track of what the automation is doing or what surrounding circumstances are. They may not understand what the implications of what they may have asked the automation to do and even forget which mode they have set the automation in. As the automation becomes more sophisticated the human's mental model may be insufficiently accurate to make the necessary inferences or predictions, such that anticipating what to do becomes difficult or impossible.

A typical solution is to say: *allocate functions to the human the tasks best suited to the human, allocating to the automation the tasks best suited to it.* This is easy to say, but not so easy to do."

At the Project it is acknowledged that the role of the operating crew is either intentionally or unintentionally created by the system design process. However, it is very little empirical data about the effects of different trade-offs of function and task allocation on the performance of control room crews. The Project are about to study the effect of different task allocation methods in particular with respect to analysing and presenting plant automatics.

A pilot experiment for HCA has recently been carried out. The main purpose was to develop and test specific performance indicators for studies of human-centred automation. Specific measures of performance are supposed to reveal the quality of performance on particular variables under

investigation. General performance measures are on the other hand valid under any operating situation, and relevant in all studies.

The specific indicators developed within the HCA-project were measures of trust, control room attribution, and self-knowledge. These measures are believed to be especially sensitive to experimental manipulations of automation, and they will supplement existing general instruments for estimation of situation awareness, workload, operator performance, and plant performance (see below).

TEST AND EVALUATION

The experimental evaluation of Computerised Operator Support Systems (COSSs) and MMIs developed at the Project has been the basis for the human factors work at Halden. The setting for this experimentation is HAMMLAB.

Over the years a well-established infra-structure and methodology for performing evaluation experiments of new operator aids and man-machine interfaces has been developed. Both operators from the Halden Reactor and operators from the Loviisa NPP in Finland take part as test subjects in these experiments. A variety of data can be collected during the experimental sessions: video and audio recordings of the activities in the control room, logs of all interactions between the operators and the simulator (displays used, control actions performed, etc.) as well as logs of critical process parameters of relevance for judging the performance of the operating crew during a particular transient. In addition, operator interviews, questionnaires, debriefing sessions, verbal protocols, etc. are utilised to extract additional information for the later analysis of the experiments.

The experiments performed are of different types. Some focus on providing direct design feedback to a specific system, while other experiments aim at providing more general knowledge for use in defining technical bases for system design and evaluation.

In addition to evaluating the final systems, there is also need for evaluations at different stages of system development to provide early feedback on the quality of certain system features. The process is therefore often iterative, running through a series of experiments on a particular system, each contributing to a better design and an improved final system.

Over the years a number of COSSs have been evaluated in HAMMLAB, (6,7). Over the last few years, work has focused on alarm systems, staffing level and night-shift work.

Alarm Systems

The need to improve alarm systems beyond the single set-point, single alarm approach has led to the development of different kinds of new alarm systems for the nuclear industry. At the Project, both a computerised alarm system toolbox - COAST (8), and a special application for HAMMLAB - CASH (9) became operational during early 1996. In 1996 CASH was used as a vehicle to carry out studies of alarm system concepts.

The purpose of the CASH evaluation program has been to test the impact of different types of alarm display, and different levels of alarm reduction, on operator and plant performance. Results from the study indicate that the operators preferred maximum alarm reduction and alarm displays, including tiles, message lists and alarms integrated into process monitoring formats.

Objective effects on operator and plant performance were reflected in interactions between display type, alarm reduction, and the availability of suppressed alarms. The main interpretations of these results are 1) the surface structure of alarm systems does not affect overall human and system performance, 2)

operators in advanced control rooms use the alarm system for limited purposes, and 3) systematic training and testing with alternative alarm systems in HAMMLAB, compensated the natural superiority of alarm systems familiar to the operators.

Staffing Level

The "Study of Control Room Crew Staffing for Advanced Passive Reactor Plants" was the largest test and evaluation project carried out by the Project in the 1994-96 time frame. It was a bilateral project done for USNRC (10).

Differences in the ways vendors expect the control room staff to interact with evolutionary or advanced plants may require reconsideration of the minimum shift staffing requirements set in today's federal regulations. This research project evaluated the impact of various level of staffing on team performance. The purpose was to contribute to the understanding of potential safety issues and to provide data to develop design review guidance.

The study was conducted both at the Loviisa NPP and at HAMMLAB. Loviisa served as the conventional plant while HAMMLAB served as the advanced plant. Data were collected from eight crews during a range of design basis scenarios, each crew serving in either a normal or minimum staffing configuration.

Results show that crews in the conventional plant experience less workload than crews in the advanced plant, and that minimum sized crew experience more workload than normal crews. The increased workload did not exceed the threshold beyond which performance degradation occurred as situation awareness, teamwork, and task performance demonstrated in the advanced plant were significantly better or higher for these measures than in the conventional plant.

Night-shift Work

Human performance is known to be greatly affected by variation in work cycles. Studies of shift-work have identified numerous incidents of peaks and troughs in human performance, in which human are at their best and worst. However, little is known how variations and work-time influences operator cognitive performance. Therefore studies of operator performance at night will be undertaken at the Project. These studies focus on the types of operator information processing and cognitive activities which are affected by the time of day. Experience gained from the experiments will constitute a basis for possible optimisation of shift-work and for modification of the man-machine interface, procedures or other relevant means.

Development of Technical Bases for Guideline Formulation

While guidelines for design and evaluation of control rooms using conventional man-machine interfaces are in general available today, the basic knowledge required for design and evaluation of computer-based control rooms needs to be further developed. Existing guidelines for computer-based systems mainly address the questions of *how* to present information (symbols, colours, font size, etc.), while little guidance is given on *which* information is important, and how process control should be executed in an efficient manner. Thus, existing guidelines for evaluation of computer-driven human-machine interfaces and digital based control rooms are incomplete. System and control room developers as well as organisations evaluating and licensing human-machine interfaces are therefore in need for guidelines in this field.

A major objective of the human factors research at the Halden Project is to provide experimental data which can contribute to formulation of guidelines for design and evaluation of computer-driven man-machine interfaces. A considerable part of the experiments in HAMMLAB are thus focusing on generic issues in connection with operator performance in an advanced control room environment. A key issue in this work is to find a suitable format for describing the generic results such that they easily can be applied by member organisations in formulation of their guidelines. Presently, the Project is preparing "lessons learned" reports from the evaluation experiments performed at Halden where emphasis is placed on making the information relevant for guideline formulation available in a structured and easily accessible way.

METHODOLOGICAL DEVELOPMENT

Eye-movement Tracking Measurements

Eye movement tracking measurement techniques have demonstrated the possibility of better studying operator cognitive activities and information processing. Studies in connection with the human error project have proved the usefulness in clarifying degraded verbalisation during period of greatest interest to researchers. A tool called EYECON has been developed by the Project to assist in analysing the enormous amount of data collected from the ETM measurements. Continued research with this technology is envisaged as part of the studies focusing on cognitive aspects of the operator.

Situation Awareness

The situation awareness measurement technique (11) is an objective measure for studying how specific support systems assist the operator in maintaining an understanding of the status and behaviour of the nuclear plant process, especially in those case where the aim is to enhance the operator's mental representation and awareness of the process. This measure has been applied successfully in a number of studies, for example in the above mentioned staffing level project. Continued research with this measure will be used to determine whether individual crew members' situation awareness are differently developed and maintained and how shared situation awareness develops and is communicated among team members.

VISA - Visual Indication of Situation Awareness

Work with a new *continuous* measures of situation awareness has recently been a central activity. The objective has been to identify advantages and disadvantages of a new measure compared to previous measures that required interruptions during the run of a scenario. The new measure takes scores from predefined areas of interest critical for solving the scenarios. The areas are scored from the eye-movement video.

Additional work is required before the measures can be validated and this will continue during upcoming HAMMLAB experiments.

Workload

There is a need to understand the effects of workload on operator performance. Studies of workload are carried out by collecting data in the control room environment, specifically the program consider the suitability of different metrics to measure subjective workload, effects of workload transition (under- and overload) on operator performance, and the interaction with and mediation of other subjective measures of operator performance (e.g. situation awareness, verbal protocols, time and accuracy analysis).

Operator Performance Assessment System (OPAS)

Systematic experiments in HAMMLAB put special demands on operator performance measurement. Performance scores have to be comparable across scenarios and sensitive to a wide range of operator competence levels. The measure should be reliable and robust, should account for team performance as well as single operator performance, and be efficient in use. It is furthermore desirable that the measure complies with established norms for human performance measurement regarding reliability, validity, sensitivity, non-intrusiveness, etc. The Project has therefore actively been developing an approach based on the prescription of optimal solutions to scenarios, based on discussion with process experts.

For each scenario, the expert solutions are represented hierarchically in a diagrammatic form. Operator activities are classified and weighted according to their importance. During the experiment, the process expert registers operator activities in real time, concurrent with operator performance. After the study a performance index is calculated estimating the discrepancy between the expert analysis and operator solutions to the scenarios.

In 1997 the Project analysed OPAS data collected during the 1996 alarm experiment and the 1997 HEAP/HCA experiment. The results show a good consistency among the activity types defined by the system, and a moderate relationship with plant performance. This relationship is strongly modified by scenario effects, but it is not yet clear which aspects of the scenario influence the degree of association, and the direction of the correlation between operator and plant performance. The OPAS measure has good variability, and only a few ceiling effects in very simple scenarios with professional operators from the Loviisa power plant have been observed. As yet no floor effects have been detected.

During 1997 an extended version of OPAS was tested for single operators. Besides objective features of operator performance (e.g. operations and detection) this version also included cognitive aspects of performance. Results so far indicate that cognitive components are highly correlated with behavioural components of the system. It should be noted that an increased number of sub-activities in the system may set the reliability and general quality of real time scoring at risk. Explicit scoring criteria for specific cognitive processes were hard to develop. These results generalise to HBWR operators with limited training in HAMMLAB, and should not impede similar studies with professional PWR operators.

Plant Performance Assessment System (PPAS)

A method for analysing plant performance measures has been developed and applied to data collected from crews participating in experiments. The performance of the crew is compared to an optimal control model developed for the scenario and afterwards a data analysis technique is applied. The applicability of PPAS is promising, and it might be an important additional measure based on real plant measures to evaluate the quality of the crew performance.

EXPERIMENTAL CONTROL ROOMS

The work on *experimental control rooms* investigates how modern technology can be taken into use to improve plant safety and efficiency. Here, experimental, but still realistic control room prototypes are built and evaluated providing useful results which can be used when upgrading existing control rooms. An important part of the programme is directed at Information Presentation Methods where issues related to navigation, process overview displays, and new forms for alarm displays are investigated.

NAVIGATION

Use of computerised systems in the CR have resulted in increasingly complex and sophisticated workstations. Typically in such environments the control system requires operators to work with computerised systems for monitoring and control of the process.

This introduces the important requirement for operators to both orientate themselves within the workspace and be able to move around in it in order to locate information. These aspects become crucial where integration of information across different displays must take place. Whilst computerisation allows presentation of greater amounts of information in more sophisticated ways than ever before, it has also increased the potential to negatively affect operator performance. This is due to the introduction of navigation problems to the system. As upgrading of traditional and hybrid control rooms occurs issues of navigation and its impact on operator performance will be increasingly important.

Navigation itself is concerned with the determination of ones position in, and the movement through, a space. This space can be a physical, a process space (mimics), or information space (network). In computerised control rooms navigation in its simplest form is related to two crucial aspects of operator performance.

- The first is maintaining or establishing awareness of the operator's position within the network of computerised displays.
- Second is the ease and efficiency with which an operator is able to travel around within the network.

At the Project a program is under way including research issues such as:

- The perceptual and cognitive aspects underpinning operators' awareness of their orientation in, and navigation through work spaces.
- Appropriate methods for representation of work space itself in order to better support operators navigating through it.
- The development of metrics for the evaluation of navigation, both in restricted experimental settings and more sophisticated simulator environments.
- Methods for evaluation and comparison of design alternatives.
- The development of guidelines on navigation for designers of HMI systems.

LARGE OVERVIEW DISPLAYS

One of the most discussed features for evolutionary and new control rooms are so-called large overview displays, one of the first examples of such a design came from Combustion Engineering through their IPSO (Integrated Process Status Overview) display. The idea is to integrate information which is normally found distributed around in the conventional control room on one permanently displayed picture where it is clearly visible to varying distances in the control room.

Project staff has designed a so-called integrated large overview display where the idea is to support rapid assessment of the plant status and dynamics by a representation of the whole process. The displays is shared by the control room staff and is a strong support for co-ordination. The layout and content of the display are context dependant to match the changing operators' needs and tasks. Centred around mimic diagrams, it combines different graphical features to support an efficient control of the complex process.

Configurable elements and an original alarm presentation support clear and rapid identification of disturbances.

ALARM DISPLAYS

CASH (Computerised Alarm System for HAMMLAB) is the newly implemented alarm system for HAMMLAB. It includes several advanced alarm handling features, such as extensive alarm structuring feasibility's and an efficient man-machine interface, aiming to reduce operator workload during plant disturbances. CASH provides two levels of presentation. At the top level, alarms are presented on the above mentioned plant-wide, high-level, system paced overview display. Here alarm information is integrated with process information rather than in a self-standing independent display.

The second level provides alarm details in detailed, totally user-controlled displays, both integrated in the NORS operating displays and in CASH selective displays. In the selective displays, which can be called up upon request from the top-field, the operator is able to look at alarms suppressed from the overview and he has access to a variety of alarm lists. By enabling a combination of the "radio-buttons" he can build the alarm list he needs for the present situation. The I/O buttons located at the bottom of the screen, enables the operator to access further information, such as trend diagrams.

PLANT SURVEILLANCE AND SUPPORT SYSTEMS

The activities on *plant surveillance and operations systems* investigates the potentials for improved plant operation through implementation of new methods and systems for plant surveillance. A common goal for both the human factors work and the systems work at the Project is to develop and test reliable and effective human-computer interfaces which can ensure operator awareness of both emerging plant situations and plant operating states. The work addresses questions related to operator tasks such as fault detection, diagnosis, prognosis and procedure implementation. One emphasis is the development and tests of actual surveillance systems, another is developing human-machine interface design proposals.

SIGNAL VALIDATION

The surveillance and control task of any industrial plant is based on readings of a set of sensors. It is essential that the output from these sensors are reliable since they provide the only objective information about the state of the process. The signal validation task confirms whether sensors are functioning properly.

A method for transient and steady state on-line signal validation has been developed at the Project using artificial neural nets and fuzzy logic pattern recognition. The method has been successfully tested on simulated scenarios covering the whole range of PWR operational conditions. Data was provided by EDF, France (12).

The neuro-fuzzy model has been implemented in a client/server software system under Windows NT. The system is called PEANO.

COMPUTERISED PROCEDURES

Procedures are important tools both operation and maintenance of a power plants. A software system for electronic handling of all types of procedures is under development at the Project. Mainstream technologies are applied in the implementation.

The procedure is represented using XML (a standard from the World Wide Web consortium). It gives support for end-user configurable procedure models. The XML format is automatically converted to HTML format. Using HTML format for procedure presentation enables use of standard web-server and web-browser products for this purpose. Stylesheets are used for tuning the procedure presentation on a per procedure model basis. Hyperlinks to other plant documents can easily be integrated.

The on-line application COPMA-III is running under the web-browser and offers functions for: login registration of end users, presenting a list of available procedures, retrieval and presentation of procedure documents with table of contents overview and functions for navigation between execution units (instructions) in the procedures.

Several users may log-in to different COPMA-III clients and co-operate in the execution of a multi-agent procedure. A procedure may also include third party software components (e.g. ActiveX or Java applets). Such components may for example provide a vendor specific interface to process equipment

ACCIDENT PREVENTION AND MANAGEMENT

The Project is carrying out a research programme on computerised accident management support (the CAMS-project). The aim is to establish a prototype of a system which can provide support to the control room operators and the staff in the Technical Support Centre during accident situations. The CAMS prototype utilises available simulator codes and the capabilities of computer-based tools to assist in identification of plant state, prediction of future development of the accident, and planning of accident mitigation strategies.

The first CAMS prototype consisted of a data base and a knowledge base, a predictive simulator and a man-machine interface system. The system was evaluated during a national emergency drill in Sweden in May 1995 with positive results. Recently new methods and modules are added for signal validation, state identification, tracking simulation, predictive simulation, risk monitoring and the man-machine interface. This second prototype is still under development. The purpose is to demonstrate that the developed functions can efficiently work together. The further plan is to test CAMS at a power plant or a national crisis centre.

ENHANCEMENT AND SAFETY ASSESSMENT OF SYSTEM QUALITY

The programme on *enhancement and assessment of system quality* addresses the issues of how to develop high quality software systems, with particular weight on methods which are relevant to safety. The work includes development and investigation of methods and tools for analysing and assessing existing software with respect to various quality aspects.

There are three complementary principles which should be followed to obtain dependable software. The first principle in this respect is fault avoidance through good software engineering and quality assurance throughout the complete life-cycle of the software. The second principle is fault detection and removal through a thorough validation and verification activity. A third principle, which could also be considered is fault tolerance, i.e. the system should be designed so that a single failure will not jeopardise safety. HRP has made research activities on methods of relevance for all these principles, as formal software development method, static analysis, testing, software diversity etc.

FAULT AVOIDANCE

There are three complementary principles which should be followed to obtain dependable software. The first principle, fault avoidance, can be obtained through good software engineering and quality assurance. However, to obtain extra high integrity the use of *Formal software development methods* has been advocated. These are methods which provide a mathematically based framework within which specification, development and verification of software systems can be done in a systematic and precise way. The use of formal specification and design makes it possible to discover many errors which might otherwise very easily be overlooked.

Following the principles behind formal software development, the Halden Project has developed a methodology based on algebraic specification and a proof tool, the *HRP Prover (13)*. One of the virtues of this methodology is that the same language, tool and proof techniques can be used both in specification and design, even down to a "concrete" specification which can be automatically translated into code. In the specification phase, the theorem prover is used to verify and validate the specification, while in the design phase the same tool is used to verify the correctness of the design steps.

There is, at the market, a variety of commercial tools supporting formal development methods. An ongoing activity at HRP is to investigate the applicability of such tools. Based on a comparative review of a large number of systems four were selected for a closer evaluation. Of these, one system has been selected for an experimental investigation, by application on a subsystem in the new experimental control room (HAMMLAB 2000) which is being developed at HRP.

FAULT DETECTION

The methods for fault detection can be divided into two main categories: *Static analysis and testing*. Static analysis is defined as the process of evaluating a computer program without executing it. The main objective of the static analysis is to check that the final program conforms with the specification or design documents, but it is also used to reveal defects in the program. These defects may be direct faults, but they may also be violation of coding standards.

The Halden Project has in the SOSAT (Software Safety Tools) project, a joint project with TÜV-Nord and GRS/ISTec in Germany developed a set of tools which can assist in the safety analysis of computer programs (14) It is based on a memory dump of the host computer, i.e. the computer where the analysed program is implemented. One reason for basing the analysis on the machine code representation of the program is to reveal potential faults introduced through the compiler and other programming aids. A disassembler extracts the part of the memory content which constitutes the program(s) and translates it into a processor independent language, which is the basis for the further analysis.

An ongoing activity at HRP on software analysis is, in co-operation with GRS/ISTec, to develop tools which check high level language programs against a variety of coding standards and guidelines.

Testing a program means to execute it with selected test data to demonstrate that it performs its task correctly. Ideally the test data should be selected so that all potentially residual faults should be revealed. The Halden Project have performed several investigations of testing methodologies. An ongoing activity at HRP on testing is an experimental evaluation of a method, the PIE (Propagation, Infection, Execution) method suggested by Dr. Jeff Voas (15).

FAULT TOLERANCE

The third principle, fault tolerance, can be obtained in different ways, e.g. through diversity and/or safety checks. A project on diverse software (PODS) was performed as a joint project between the Safety and Reliability Directorate (SRD), Central Electricity Research Laboratory (CERL), The Technical Research Centre (VTT) of Finland and HRP (16). The main objective of the project was to provide a measure of the relative merits of using diverse programs, as compared with any one of the programs replicated in all channels, in a 2-out-of-3 majority voting protection system. To achieve the objective, an experiment was mounted which simulated a normal software development process to produce three diverse programs to the same requirement. The requirement was for a reactor over-power protection system. After careful independent development and testing the three programs were tested back-to-back against each other to locate residual faults. The three development teams were allowed to discuss problems with the requirement specifiers, but not with each other. All phases of the project were carefully documented for subsequent analysis.

FURTHER PLANS

THE HAMMLAB 2000 PROJECT

The Halden Project has experienced an increased demand for a facility able to support advanced human factor's related experiments, both through members organisations in the joint research programmes, as well as through requests for doing specific studies for certain organisations on a bilateral basis. The Project therefore set forth to investigate whether today's HAMMLAB would be able to meet tomorrow's needs for an experimental facility, and the HAMMLAB 2000 project was initiated. The HAMMLAB 2000 project has the goal of establishing a flexible infrastructure regarding the physical laboratories and the hardware and software, as well as making sure tomorrow's HAMMLAB has a broad pool of simulators. By the year 2000 four simulators will be available providing broader scope and details in the simulation and access to a more operators to participate in experiments.

The NORS Simulator

The NORS simulator is a "westernised" VVER simulator of the Loviisa nuclear power plant in Finland. NORS was manufactured in 1983, and has since then been the nucleus of HAMMLAB. Several modifications and additions have taken place of the NORS simulator models, and NORS is now considered to be a very good simulator for the purpose of being "the process" when performing experimental studies. NORS will continue to play a role also in the coming years as part of the HAMMLAB 2000 pool of simulators.

The CP0 Fessenheim Simulator

The member organisations of the Halden project have clearly expressed the wish for a western type of PWR simulator as part of the HAMMLAB 2000 pool of simulators. In co-operation with Thomson Training & Simulation and Electricite de France a delivery of a full-scope simulator of the Fessenheim-1 plant in France took place in 1998. Fessenheim-1 is a Westinghouse-like 900 MW 3-loop plant built by Framatome.

The Forsmark-3 BWR Simulator

In order to prepare for maximum transferability of results from experimental studies, it has also been decided to include a simulator of a BWR plant in HAMMLAB 2000. BWR utilities in Sweden and

Finland have stated their interest in having a BWR simulator as part of HAMMLAB 2000, since they see a clear benefit to be able to run specific studies in HAMMLAB to feed results into their large control room modernisation programs.

It has been agreed between the Swedish and Finnish utilities and the Halden Project to make a BWR simulator based upon the Forsmark 3 NPP, a 1160 MW BWR located north of Stockholm, Sweden. The simulator is manufactured by VTT Energy in Finland in co-operation with the Halden Project, and is scheduled for completion in the spring 1999.

The Oseberg Simulator

The Oseberg training simulator is a full-scope simulator of the Oseberg A oil production platform located in the North Sea. It will be used as "the process" when performing human factors studies related to the oil and gas industry. The Oseberg simulator has been ported to modern hardware and will be modified to handle the recent and future developments in the North Sea, e.g. operation of satellite fields from a centralised control room and remote control of production onshore.

SUMMARY

Backfitting of nuclear power plant control rooms is a continuing process, introducing computer-based solutions for surveillance and control as well as for improving the human-computer interface. At the same time designs for tomorrow's reactors are developed, characterised by fully digital instrumentation and control systems, and advanced, computer-based control rooms. Research and development efforts are needed to ensure that the new technology gives the expected improvements in operational safety and efficiency.

The research programme at the Halden Project addresses the research needs of the nuclear industry in connection with introduction of digital I&C systems in NPPs. The programme provides information supporting design and licensing of upgraded, computer-based control room systems, and demonstrates the benefits of such systems through experiments in its simulator-based experimental control room facility at Halden.

At the Halden Project an internationally sponsored research programme is carried out which addresses these research issues. The Halden Man-Machine Laboratory represents a unique test-bed for investigating new, computer-based solutions for nuclear power plant control rooms. The research programme draws upon competence built up through more than 25 years work in the field of computer-based operator support and digital control room solutions, and the close contact with licensing authorities, utilities and reactor vendors in the 20 countries participating in the Halden Project ensures that the work is addressing the real research needs of the nuclear industry.

With the initiation of the HAMMLAB 2000 project the goal is to broaden the scope and domain of the human factors studies, by introducing two new nuclear process simulators and one petroleum simulator. It will be possible to perform studies related to severe accidents due to the extended operational domain of the new simulators.

HAMMLAB will be expanded with new laboratory areas making it possible to perform experiments in parallel, without interfering each other. The laboratories will be flexible with regards to physical size, in order to adapt to small and large studies. A software and hardware infrastructure will be developed, allowing for integration of software systems developed at the Halden Project or in member organisations. Such systems can then be tested in a realistic environment, prior to installation in real-life plants.

Based on the above, the Halden Project is confident that the new HAMMLAB will be the global centre of excellence for performing human-machine interaction studies for the management and control of industrial processes, when it is taken into use in the year 2000.

REFERENCES

1. Førdestrømmen N.T., Kvaalem J.: *Halden Man-Machine Laboratory as of 1996*, Halden Work Report, HWR-476, May 1996.
2. C. Manetta and R. Blade in "Glossary of Virtual Reality Terminology" in the International Journal of Virtual Reality, Vol.1 Nr.2 1995.
3. C.Holmstrøm, M.Louka, F.Øwre "Human Factors Engineering and Control Room Design. Using a Virtual Reality based Tool for Design, Test and Training." Water Reactor Safety Meeting, Washington, USA. October 1998.
4. M Green, E. Hollnagel et al: "Overview and Results from the Human Error Analysis Project 1997-1998" Water Reactor Safety Meeting, Washington, USA. October 1998.
5. T.Sheridan: "Human centred automation - A tutorial" EHPG Loen May 1996.
6. E.C.Marshall et al: "The experimental evaluation of the Success Path Monitoring System" IEEE Fourth conference on Human factors and Power plants, Monterey June 1988.
7. C.B.O Holmstrøm et al.: "Continued experimental evaluations of a diagnostic rule-based expert system for the nuclear industry" ANP'92 international conference on design and safety of advanced NPPs. Tokyo, October 1992.
8. A.Bye et al: *COAST - Alarm System Toolbox*". 2nd IFAC workshop on Computer systems and AI in Process control. Lund, Sweden August 1994.
9. B.Moum et al: "CASH- the new alarm system for HAMMLAB" HWR-480. EHPG Loen May 1996.
10. B. Hallbert, A. Seebok et al: "Interim results of the Study of control room crew staffing for advanced control rooms. Water Reactor Safety Meeting, Washington, USA. October 1996.
11. D. Hogg et al.: *Measurement of the operator situation awareness*" HWR-377, September 1994.
12. P. Fantoni, F. Øwre: *Full range signal validation of PWR data* Water Reactor Safety Meeting, Washington, USA. October 1998.
13. T. Sivertsen: *Putting principles into practise - the formal development of a Theorem Prover*" Water Reactor Safety Meeting, Washington, USA. October 1998.
14. G. Dahll M. Barnes, P. Bishop: "Software Diversity: way to Enhance Safety." Information and Software Technology, vol. 32 no. 10, 1990.
15. J.M. Voas: "PIE: A Dynamic Failure-Based Technique," IEEE Trans. Software Eng., 18 (8), 717-727, 1992
16. G.Dahll and J.E. Sjøberg: "SOSAT - a Set of Tools for Software Safety Assessment" Paper presented at Second European Conference on Software Quality Assurance, Oslo 1990.

Overview and Results from the Human Error

Analysis Project 1997-1998

Per Øivind Braarud, Asgeir Drøivoldsmo, Erik Hollnagel, Mark Green
OECD Halden Reactor Project, P. O. Box 173, N-1751 Halden, Norway
Email: <firstname.lastname>@hrp.no

ABSTRACT

The ongoing Human Error Analysis Project (HEAP) was initiated within the Halden Reactor Project in 1994. Its objectives are to develop a better understanding and explicit model of how and why 'cognitive errors' occur, and to provide design guidance to avoid, or compensate for, cognitive errors. During 1994-1996, results led to practical insights concerning diagnostic strategies and styles, aspects of operator support systems affecting diagnosis, and success rates for diagnosis and error types. From 1996 the project's scope was extended to consider error prediction, error recovery, method development, and investigation of complexity.

HEAP has investigated human error within complex simulator-based experiments using the Cognitive Reliability and Error Analysis Method (CREAM), and preliminary results are presented. Extensive method development has also been carried out, to allow the study of operator cognitive activity within realistic situations. New measures for operator performance, situational awareness, and plant performance, have been developed and are briefly described. Understanding of what makes a control room situation difficult to handle is important when studying operator performance, with respect to both prediction, and improvement of the human performance. Therefore, HEAP has been investigating the complexity of the operator's work situation. From these investigations a definition and measure of complexity are being developed. A Complexity Profiling Questionnaire has been developed, based on factor analytic results from operators' conception of complexity. Initial validity of a set of identified complexity factors has been shown, by prediction of both crew and plant performance from ratings of the complexity of scenarios.

Human Error Analysis Project (HEAP)

The OECD Halden Reactor Project has, since 1994, been engaged in a long term effort to study human erroneous actions. The purposes of this study are: (1) to provide a better understanding and explicit modelling of how and why erroneous actions occur (specifically when they involve cognitive activities such as diagnosis), and (2) to provide improved design guidance for the development of man-machine systems, that can avoid or compensate for erroneous actions.

Four initial pilot studies put the emphasis on methodological aspects, in particular the development of methods to investigate cognitive aspects of behaviour using real operators in difficult scenarios (Kaarstad et al., 1994 & 1995; Kirwan, 1994a).

The objective of the first pilot study was to evaluate concurrent and interrupted verbal protocol techniques, with regard to their applicability to identify operators' problem solving strategies, diagnostic strategy types, and possible "cognitive inefficiencies" while locating a fault. The main finding was that the two methods provided different types of information about problem solving strategies and human erroneous actions. The methodological recommendation for the main HEAP study was, therefore, to use a combination of the two methods.

The objective of the second pilot study was to test whether eye movement tracking analysis was a feasible supplement to verbal protocols in studying operators' cognition during fault-finding. The main finding was that the analysis of eye movements helped to make the interpretation of verbal protocols more robust and gave a better insight into the operators' problem-solving behaviour. Wearing the eye-tracking equipment had no apparent effect on the quality or quantity of verbal protocols, or upon diagnostic performance.

The third pilot study investigated the effects of scenario complexity on operators' diagnostic behaviour. Complexity was assumed to be a multidimensional concept that was varied by manipulating the number of underlying faults in three different scenarios. The main finding was that the number of underlying faults did not, by itself, prove to be a dominant complexity factor, when the performance measures were the degree of operator success in diagnosing the faults, and the use of different diagnostic strategies. This study also used operators from two different operating environments but the results showed no systematic variation with respect to either subject pool, or performance measures.

The fourth pilot study looked at the quality of information provided by different data sources. The preceding studies had shown the need for multiple data sources, typically concurrent plus interrupted verbal protocols and eye movement data. Since considerable resources are needed to analyse a combination of different data sources, it is important to know in advance what the relative contribution is from each source of data. The findings were that all three types of protocols (concurrent, auto-confrontation, and expert) produced similar results for a set of pre-defined target activities, although concurrent verbal protocols were the richest source of data. Furthermore, it was found that concurrent verbal protocols can effectively be used for teams as well as for single operators.

Performance Prediction In HEAP

The pilot experiments had demonstrated how diagnostic errors could be found from performance records (Follesø et al., 1996), but another essential issue is to be able to predict the likely error modes. In order to achieve this an experiment was designed with the purpose of developing and refining a method, or set of methods, for predicting performance failures, specifically the error modes that can be expected for a specific task. The method should specifically enable the analysts to:

- identify the types of incorrect performance (error modes, cognitive failure modes) that are possible for the given task or scenario;

- qualitatively rank or rate the possibility of these possible error modes in order to identify those that are the more likely to happen.

The basis for any kind of performance prediction must be a detailed description of the situation where the performance takes place and a specification of the critical aspects of the performance. Since the error modes are basically deviations from the expected performance, the starting point for making a prediction of this type must refer to a description of the expected performance. This can either be developed from scratch or make use of already existing descriptions from, e.g., a task analysis, existing operating procedures, ideal paths or performance time-lines and event trees such as those used in PSA/HRA.

In this case, the predictions were based on the task descriptions provided by the Operator Performance Assessment System (OPAS), see below for description. For each scenario OPAS provided a description of the operators' actions, using a simple hierarchy of main goals, subgoals, and operator tasks, where the latter were classified as either detection or operation.

Performance Prediction Method (CREAM)

The basic steps of the prediction method used in the experiment are shown in Figure 1. The method was derived from the basic method of CREAM (Cognitive Reliability and Error Analysis Method), cf. Hollnagel, 1997.

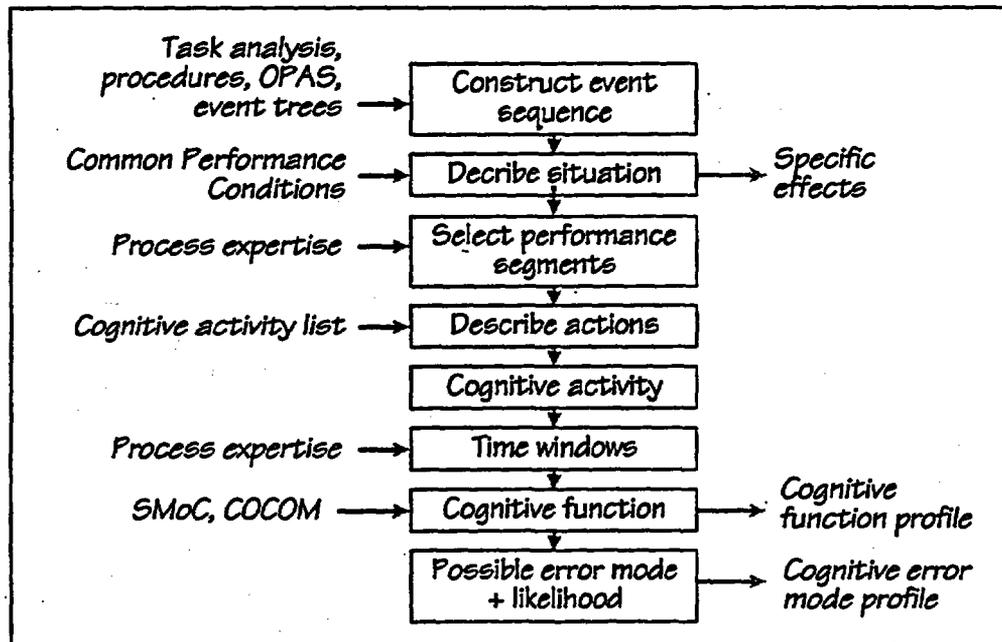


Figure 1: Basic steps of performance prediction.

The actual experiment was carried out together with the NRC Alarm Project which aimed to investigate the effect of various types of alarm display and alarm processing methods. The experimental design is summarised in Table 1. The main independent variables in the alarm experiment are three kinds of alarm display types and three alarm processing levels. However, the experiment was carried out in only eight experimental conditions because of limited experimental resources such as subject schedule, time requirement, and so on. The selected eight experimental conditions are also shown in Table 1.

Table 1. Experimental Conditions in the Alarm Experiment

| Display type | ALARM PROCESSING LEVEL | | | | |
|--------------|------------------------|--------------------|--------------|---------------------|--------------|
| | No processing | No nuisance alarms | | No redundant alarms | |
| | | No availability | Availability | No availability | Availability |
| Tile | 1 | 7 | | | |
| Mixed | 2 | 3 | 4 | 5 | 6 |
| Integrated | | 8 | | | |

The HEAP experiment used the conditions where performance differences were thought most likely, i.e., conditions 2 (no processing, mixed display) and 8 (nuisance alarms removed, integrated display). In addition, condition 7 (nuisance alarms removed, tile display) was selected as a base line condition for comparison purposes. This selection resulted in a total of 36 scenarios to be analysed.

The experiment was conducted in the Halden Man-Machine LABORatory (HAMMLAB) at the Halden Reactor Project using a full-scope simulation of a PWR NPP. The subjects were 12 licensed commercial power plant operators (six crews of operators participated with two operators per crew) from the Loviisa NPP in Finland.

The data includes a record of the alarms presented to the operator during the scenario, all the operators' interactions with the simulator, and the manipulations performed by the experimental leader. The variable log includes the process parameters known to change during the scenario. Other data sources were eye movements and video recordings, together with soundtracks of the operators' communications and those of an expert commentator. Furthermore, the OPAS checklist was used to score whether the operators followed the predefined operating sequence. After the experiment this checklist could be compared with the simulation log and video recording to validate the scoring and calculate performance rating.

OPAS was used as a main source for both the predictions of error modes and for analyses of the observed errors during the experiment. The predictions were analysed using the time windows, the process expert rated OPAS formats, simulator data and expert commentators. Video and eye movement data were used for clarifications when needed.

Results

To evaluate the precision of the predictions, the scenarios were scored independently by two analysts. The agreement was high, with a mean of about 72%, ranging from 53% to 88% agreement. The quality of the interpretation can be determined by using a test of interscorer reliability, such as Cohens Kappa (Breakwell et al., 1995). Cohens Kappa compares the nominal scales from the scoring and takes into account the probability of the same scoring due to chance. The value of Cohens Kappa was calculated to be 0.66 which is significant ($p < .05$).

All predictions were analysed and a match percentage for each scenario was calculated. The match percentage ranged from 42% to 100%, with an mean of 67.8%. The histogram (Figure 2) shows the hit percentage distribution of the different scenarios.

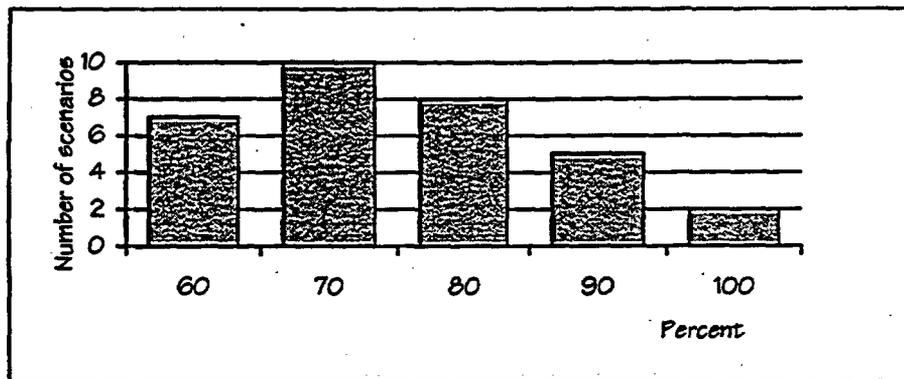


Figure 2: Percentage match of scenario predictions.

A more detailed analysis showed that the overall match between predicted and observed observation error modes was 100%, while there were no observed planning errors. Interpretation errors had a 68% match and execution errors had a 56% match.

The predictions also included an evaluation of what the most likely error mode to happen would be for each scenario. This was based on the scorers' judgements together with the characterisation made by the common performance conditions of the scenario. The error mode judged to be the most likely for each scenario, actually occurred in 72% of the cases.

Not all error modes were predicted. This is probably due to the fact that OPAS was made for a different purpose and that the task analysis therefore served a different objective. If more work is put into defining the operator tasks, a higher match between predicted and actual error modes is expected. This will be investigated in the near future and reported during 1999.

Method development related to HEAP since 1996

The first three years of the HEAP project also revealed a need for better performance measurement. Measures of human task performance are typically developed and applied in

limited laboratory settings, where the participants carry out unrealistic and well-defined tasks. The performance indicator is usually accepted by convention, and without discussion. Unfortunately, a direct transfer of such methodologies to handling complex problem solving in dynamic operating environments, which is the essence of the experimental activities in HEAP, produces instruments that are inflexible and unable to meet the demands of validity, and experimental control.

The work within HEAP has been based on experiments with a high degree of ecological validity. Therefore, in order to maintain a realistic experimental situation a requirement for the methods used has been that they must not disturb or influence operators problem solving. This requirement has led to the development of a new set of measures capturing the different categories of performance involved in simulator transients with a risk of human error. An assumption has been that there is a causal relationship between plant behaviour, operator actions causing the plant behaviour, and the operators cognitive processes causing the operator actions.

These three categories have been investigated using three different measures; plant performance, operator performance, and the operators' cognitive processes. *Plant performance* puts focus on the outcome of operator problem solving in the control. *Operator performance* refers to a direct evaluation of the quality of operator activities. *Cognitive processes* relates to cognitive predictors of system control, e.g., situation awareness. The relation between these categories and the developed measures are shown in Figure 3.

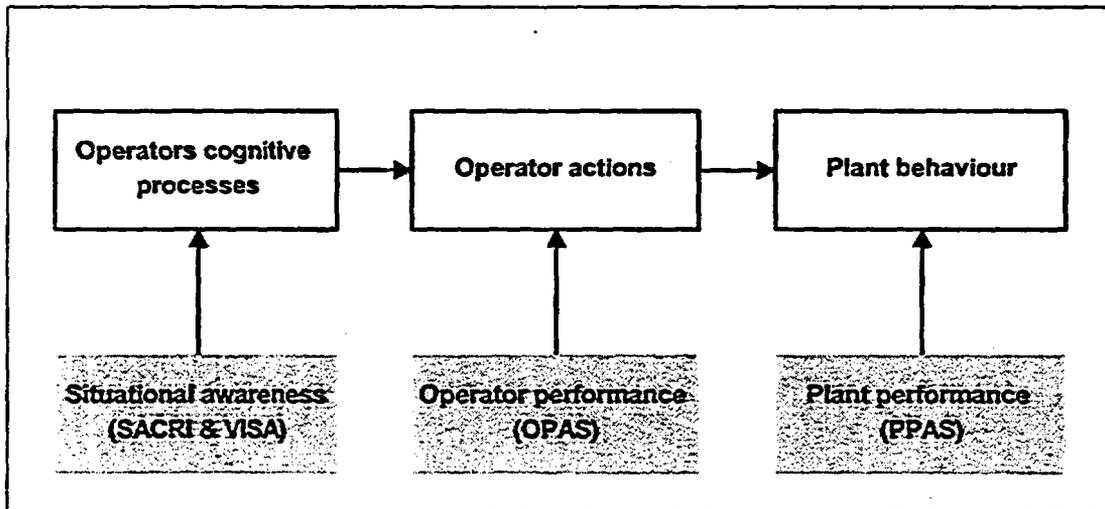


Figure 3: Relationship between the measures developed at Halden and performance categories

Plant performance

Plant performance measures representative of a crew's control of the plant, are especially useful in assessing the global performance of the crew in the scenario. During the simulation of a transient, such as a small loss of coolant in a pressurised water reactor (PWR), the minimum level reached in the pressuriser is a measure of the effectiveness of the operator corrective actions, independent of their specific strategy for tackling the problem. Such a measure, depending on the objectives of the concept under test, might be more significant than a specific action, e.g., pipe isolation, because it provides an integrated measure of the effectiveness of the operator control of the plant.

Because process control measures are taken at the end of the sequence of general operator tasks, i.e., state identification, action planning, and action implementation, they provide an integrated assessment of the operators impact on the plant rather than focusing on any particular sub-task. Additionally, these measures are meaningful because they can be directly related to the production and safety goals of the plant.

The procedure for applying the plant performance measures involves the following. For each critical process parameter in the scenario, the relative error between the crew's value and the optimal value is calculated. The optimal value is identified from the same scenario run by an expert. The relative error is the absolute difference between the crew's value for the process parameter and the optimal value, divided by the optimal value. The score of the crew in the scenario is the relative error of each parameter multiplied by the correspondent weight for that parameter, and summed for all parameters in the set.

Using this approach, good results were found in an experimental evaluation study of alarm systems, conducted in HAMMLAB. The results are reported in Moracho (1998).

Operator performance

The need for a special instrument handling operator performance resulted in the development of OPAS, the Operator Performance Assessment System (Skraaning 1998). This system prescribes optimal solutions to scenarios through discussion with process experts. For every scenario, expert solutions are represented in diagrams in a hierarchically manner. During the actual experiment, a process expert registers operator activities in real time, concurrent with operator performance. Performance scores are then calculated, estimating the discrepancy between the expert analysis and operator solutions to scenarios.

Using OPAS, performance is comparable across scenarios, the system can account for individual solution paths in dynamic operating environments, and it can be applied to crews or single operators without adjustment. Real-time registration of operator activities has been found to be very effective, and the performance assessment can be accomplished using reasonable time and resources. An empirical evaluation detailed in Skraaning 1998, indicated that OPAS complies with most established standards for human performance measurement.

Situation Awareness

Based on techniques from the aviation area, the Situation Awareness Control Room Inventory was developed some years ago at Halden (Hogg et al. 1995). One drawback of this method however is its lack of flexibility as questionnaire based data collection must be done at regular intervals throughout the scenario. This process therefore interrupts the scenarios and may interfere with some aspects of operator decision making, as well as being inflexible because the time windows for applying SACRI must be decided before hand.

Preliminary testing of a continuous measure of situation awareness is reported by Drøivoldsmo et al. 1998. The report presents a method for continuous measures of situation awareness, based on analyses of the operator's eye movements. The objective is to identify, develop, and test the new measure, and compare it to instruments that require interruptions of scenarios. Using experimental data from the 1996 CASH/ NRC Alarm study and the 1997 Human Error Analysis Project/ Human-Centred Automation study, the new measurement techniques have been tested and evaluated on a preliminary basis. The results showed promising relationships between the new continuous measure of situation awareness, and established instruments based upon scenario interruptions.

Complexity research in HEAP

Understanding of what makes a control room situation difficult to handle, is important when studying operator performance with respect to both prediction, and improvement of the human performance. The fact that control room work is complex, especially for non-normal situations, is emphasised by both practitioners and researchers. To address this area HEAP initiated work in order to understand what makes scenarios complex and to identify those factors associated with diagnostic challenges. Important questions are: What is the basic complexity factors of the work situation, and how can those factors be measured? What complexity factors are related to what types, and level, of human performance?

A basic distinction can be made between 'subjective' and 'inter-subjective' complexity. An operator carrying out a scenario has a subjective experience of the task and the situation. The operator's perception of the complexity can be labelled 'subjective complexity'. Alternatively the operator's work situation for a scenario can be evaluated for complexity by persons with substantial operational knowledge. This kind of rating is like asking 'how complex is this scenario for a typical operator'. This evaluation can be labelled 'inter-subjective' complexity. Determination of inter-subjective complexity for a work situation will typically involve operational experts' evaluation. A measure or description of inter-subjective complexity is often based on some kind of analysis of the situation, as opposed to subjective complexity that involves dynamic experience.

Complexity is often seen as related to performance. The goal of this research is to discover relations between complexity factors and human performance. A description and a measure of

the complexity of the operator's work situation should relate to operator performance. This is the case for both expert-rated, inter-subjective complexity and subjective complexity.

Figure 4 shows the work on complexity that was undertaken in 1995 and 1996. The results from this work have been further analysed in 1997 and 1998.

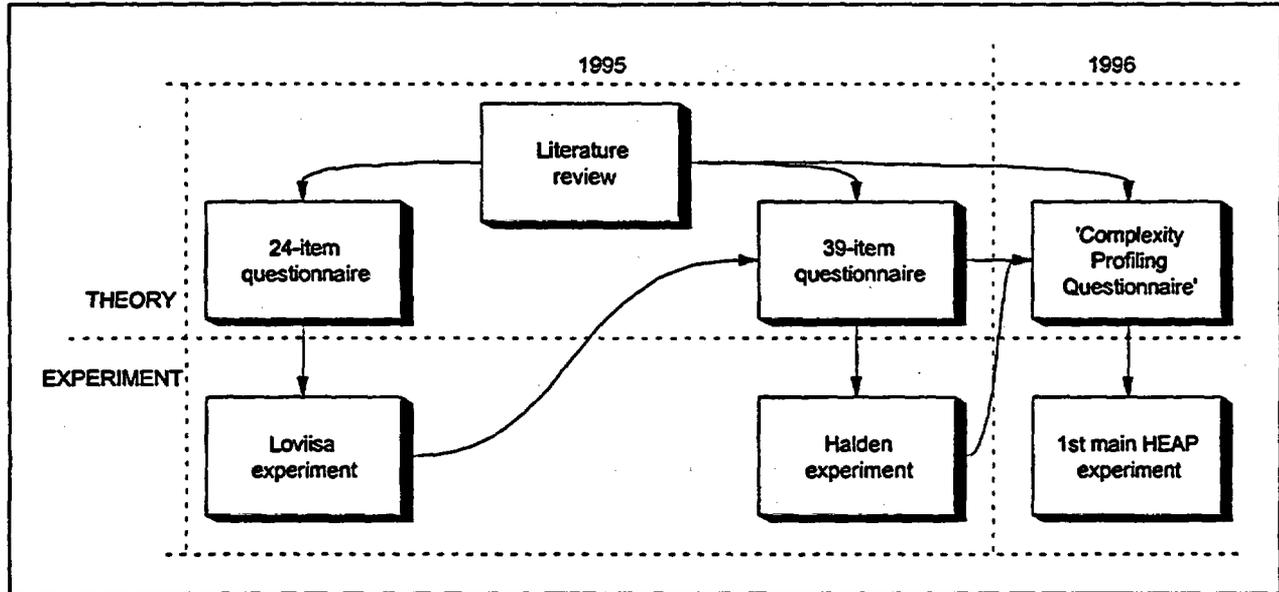


Figure 4. Work on Complexity at Halden Reactor Project

A literature review was carried out in 1995 for identification of complexity components. Based on this review a questionnaire for rating of complexity was developed, tested in one experiment, and further developed to a 39-item questionnaire (see Collier, 1998). The 39 item questionnaire was administered to operators participating in an experiment in 1995 in HAMMLAB. For a further description of the study see Hallbert, Sebok and Morriseau (1997). Twelve licensed operators, from the Loviisa NPP in Finland, participated in the study. The 12 operators made up four crew configurations with positions as shift supervisor, reactor operator, turbine operator and balance of plant operator. All operators participated in five scenarios. After each scenario, each operator completed the 39 item complexity questionnaire, giving 12 ratings for each of the 5 scenarios producing a total of 60 ratings.

A factor analysis of the operators' answers to the 39 questions identified 8 factors. The 8 factors and their interpretation are given in Table 2 below.

Table 2: The eight factors identified, with brief description

| <i>Factor</i> | <i>Description</i> |
|--|---|
| Root cause difficulties | Symptoms of the disturbance are masked for the operator. |
| Spread of information | Spread of information on the Human-Machine Interface. |
| Confusion | Ambiguous or misleading information. |
| Breadth of information gathering and co-ordination | Attention to other persons work and information from different systems. |
| Obviousness | No clear information pointing to the fault. |
| Attentional demand | Information load, including information load from alarms. |
| Severity | Challenge the safety of the plant and require fast action. |
| Temporal demand | Time pressure and many simultaneous tasks. |

These factors can be seen as a description and definition of one aspect of the complexity of the work situation in a NPP control room. The factors identified served as a basis for the development of a Complexity Profiling Questionnaire, see Braarud, 1998 for a full description.

Complexity is often seen as inversely related to performance. An inverse relation between the complexity measure and performance demonstrates validity of the complexity measure. In the second experiment prediction of performance was studied. Figure 5 outlines the predictions studied.

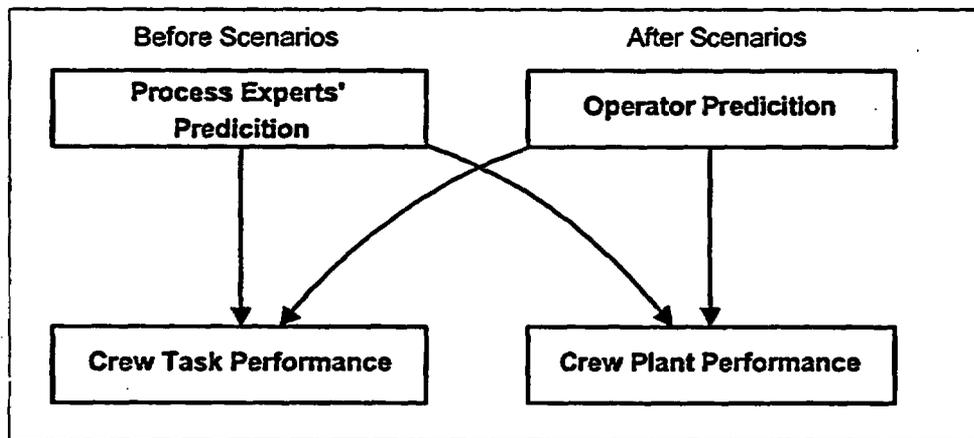


Figure 5. Prediction of performance from complexity ratings.

In the second experiment both operational experts' and the operators' running the scenario rated the complexity by use of the CPQ. Operational experts rated the scenarios before the experiment was carried out, and the operators rated the scenario after they had finished the scenario. In this experiment twelve professional operators made up six crews with one reactor operator and one turbine operator in each crew. Sixteen scenarios, that represented both high and low complexity situations, were run through by all six crews. For a complete description of the method in the study see O'Hara et. al (1997).

The operational expert's rating of the eight CPQ dimensions was used to predict the crew's task and plant performance. The same was the case for the operator's subjective complexity. The measure for crew task performance was the Operator Performance Assessment System (Skraaning, 1998), and the measure for crew plant performance was the Plant Performance Assessment System (Moracho, 1998). The prediction was done by multiple regression where performance was the dependent variable, and the ratings of the 8 complexity factors were the independent variables, see Braarud, 1998. The results from the predictions were:

- Experts' ratings predict Crew Performance. ($R^1 = .64, p < .001$)
- Experts' ratings predict Plant Performance. ($R = .69, p < .001$)
- Operators' ratings do not predict Crew Performance. ($R = .37, p = .09$)
- Operators' ratings predict Plant Performance ($R = .62, p < .001$)

The results from the multiple regressions above suggest that experts' rating of complexity, by use of the complexity profiling questionnaire, predict both crew performance and plant performance quite well. The results further suggest that operator's rating of complexity, by use of the same questionnaire, predicts plant performance quite well. This finding gives some initial validation of the complexity dimensions identified and the developed Complexity Profiling Questionnaire.

Crew performance was not predicted by the operators' ratings. Crew performance and plant performance can be seen as different aspects of performance according to the crew's goals. The plant performance measure is more directed towards optimal physical performance of the components of the plant, while the measure of crew performance is directed more towards the security of the plant as manifested in the control room crew's responsibilities.

Together these findings suggest a good potential for further and more detailed research on complexity. Further research concerning the structure of complexity is needed, not least to give a better foundation for the detailed study of which complexity shaping factors are related to what kind of scenario characteristics, and to what kind of operator behaviour.

¹ R = multiple correlation

Conclusion

Halden Reactor Project has, since 1994, studied operator performance via the Human Error Analysis Project, the objectives being to develop a better understanding and explicit model of how and why 'cognitive errors' occur, and to provide design guidance to avoid, or compensate for, cognitive errors. Following the first three year period, 1994-1996, the project's scope was considerably extended to consider error prediction, error recovery, method development, and investigation of complexity.

These difficult issues have been investigated within complex, simulator-based experiments using a wide variety of methods. Work on error prediction has been carried out using the Cognitive Reliability and Error Analysis Method (CREAM), and preliminary results are promising. The difficulty of studying operator performance in realistic situations has also placed heavy requirements for the development of new and better methods and measures. Therefore, extensive method development has also been carried out within HEAP, to allow the study of operator performance, situational awareness, and plant performance.

Important work in the area of complexity has also been carried out, and this is starting to provide an understanding of those factors that make a control room situation difficult to handle, and is especially important when studying operator performance with respect to both prediction, and improvement of the human performance. The Complexity Profiling Questionnaire has been developed, and the validity of the set of complexity factors has been shown by prediction of both crew and plant performance from ratings of the complexity of scenarios.

During 1999, the end of the second three year period, HEAP will continue to address this broad range of issues. It is planned not only to conduct further experiments, but also to produce a number of reports summarising the work done, and lessons learned to date. The opportunity to further analyse existing data in order to both clarify issues, and explore new ones, will be taken.

References

- Braarud, P.Ø. (1998) *Complexity Factors and Prediction of Crew Performance*. OECD Halden Reactor Project, Halden Work Report-521, Halden, Norway.
- Breakwell, G.M., Hammond, S & Fife-Schaw, C. (Eds.) (1995). *Research Methods in Psychology*. London: SAGE Publications.
- Collier, S. G. (1998). *Development of a Diagnostic Complexity Questionnaire*. OECD Halden Reactor Project, Halden Work Report-536, Halden, Norway.
- Drøivoldsmo, A., Skråning, G., Sverbo, M., Dalen, J., Grimstad, T. & Andresen, G. (1998). *Continuous Measures of Situation Awareness and Workload*. OECD Halden Reactor Project, HWR-539, Halden, Norway.

- Follesø, K., Kaarstad, M., Drøivoldsmo, A., Hollnagel, E. & Kirwan, B. (1996). *Practical insights from initial studies related to the human error analysis project (HEAP) OECD Halden Reactor Project (HWR-459)*. Halden, Norway:
- Hallbert, B. P., Sebok, A. L., and Morisseau, D.(1997). *A Study of Control Room Staffing Levels for Advanced Reactors (Draft)*. U. S. Nuclear Regulatory Commission, NUREG/IA-0137.
- Hogg, D. N., Follesø, K., Strand-Volden, F. & Torralba, B. (1995). *Measurement of the operator's situation awareness for use within process control research: four methodological studies*. OECD Halden Reactor Project, HWR-377, Halden, Norway.
- Hollnagel, E. (1993). *Human reliability analysis: Context and control*. London: Academic Press.
- Hollnagel, E. (1997). *Cognitive reliability and error analysis method*. London: Elsevier.
- Kaarstad, M., Follesø, K., Collier, S., Hauland, G. & Kirwan, B. (1995). *Human error - the second pilot study (HWR-421)*. OECD Halden Reactor Project, Halden, Norway.
- Kaarstad, M., Kirwan, B., Follesø, K., Endestad, T. & Torralba, B. (1994). *Human error - the first pilot study (HWR-417)*. OECD Halden Reactor Project, Halden, Norway.
- Kirwan, B. (1994a). *Human error project experimental programme (HWR-378)* OECD Halden Reactor Project, Halden, Norway.
- Moracho, M. J. (1998). *Plant Performance Assessment System (PPAS) for crew performance evaluations. Lessons learned from an alarm study conducted in HAMMLAB*. OECD Halden Reactor Project, HWR-504, Halden, Norway.
- O'Hara, J. M., Brown, W. S., Hallbert, B., Skråning, G., Peresensky, J. J., and Wachtel, J. (1997). *The Effects of Alarm Processing and Display on Operator and Plant Performance (Draft)*. U.S. Nuclear Regulatory Commission.
- Skråning, G. (1998). *The Operator Performance Assessment System (OPAS)*. OECD Halden Reactor Project, HWR-538, Halden, Norway.

Prospects on Combining Software QA Techniques

Terje Sivertsen

OECD Halden Reactor Project

P.O. Box 173, N-1751 Halden, Norway

Abstract

On basis of several research activities at the OECD Halden Reactor Project, this paper discusses some of the options available for combining software quality assurance techniques. While the various options reflect a wide variety of techniques, several classes of combinations involve the use of formal methods. In particular, the combination of graphical and textual notations represents a promising approach to making formal specifications comprehensible to a wider group of users. The paper presents research results related to the combination of Petri nets and algebraic specifications, as well as to the development of a graphical front-end to the editing of algebraic specifications. The practicality of formal methods is further enhanced by a proper combination of formal methods with CASE-tools and traditional development techniques. The paper presents results from a research project concerned with using, and measuring the effect of, formal methods in real-life software development. A third class of combinations concerns the use of complementary formal verification techniques. A distinction is made between theorem proving and model-checking, both of which represent well established approaches to the verification of software specifications. On basis of the advantages and shortcomings of the respective techniques, consideration is given to the possibilities of a combined approach. There are also several options available for the combination of program testing techniques. While formal verification is concerned about proving correctness of a specification, testing normally focuses on the (possibly symbolic) execution of a specification or a program. The paper discusses how testing can be made more focused if it is guided by the PIE-technique, which is a dynamic failure-based technique for performing program sensitivity and testability analysis. Finally, the paper demonstrates how the integration of software process- and product quality relates to the need to incorporate a variety of factors into quality assurance or assessment. It is demonstrated how these factors, which represent disparate evidences of software quality, can be combined in quality assessment by means of so-called Bayesian networks.

1 Introduction

The OECD Halden Reactor Project has for more than 20 years been working actively within the field of software verification and validation. The importance of these activities relates to the high requirements put on embedded software in computer based systems for the control and supervision of nuclear power plants.

This is reflected in the Halden Project research programme, which emphasises methods for improving software quality, and for assessing programmable systems with respect to quality attributes such as correctness, safety, and maintainability.

In many application areas, the specification and analysis of a software system is accomplished using a wide variety of notations and techniques. The lack of coherence between notations often represents a serious problem in the verification and validation of the resulting system. This is a particular problem in the translation of a requirements specification, often written by a plant engineer or other type of process specialist, into a design specification that constitutes the document which a computer specialist uses as a basis for the development of a computer based system. Experience shows that many program defects are caused by errors made in this phase of the development. This indicates that the effectiveness of software quality assurance largely depends on the success of combining complementary notations and techniques. While this is a frequent observation both in research and practise, there is still a lack of consensus on the usefulness of the various combinations. In many instances, even the effectiveness of a single technique is difficult to measure. The Halden Project addresses these problems by investigating how different notations and techniques can be combined in order to improve the overall scope and effectiveness of software specification, verification, and quality assurance.

The contents of this paper is based on work performed by several research scientists at the Halden Project. Section 2 discusses the problem of facilitating the use of textual formal notations by the complementary use of graphical techniques. The use of formal methods in combination with CASE tools and more traditional techniques is discussed in section 3, with particular emphasis on the possibilities of measuring the effect of formal methods. Section 4 discusses theorem proving and model checking as two complementary approaches to formal verification. Program testing is covered by section 5, where emphasis is given to the combination of software sensitivity analysis and reliability assessment. Finally, section 6 discusses the problem of integrating software process- and product quality, and how it can be approached by the use of so-called Bayesian networks.

2 Graphical and Textual Notations

A common opinion in industry is that both formal and conventional software specifications are ineffective in reducing the amount of software errors because they are not comprehensible to the process engineers responsible for the system requirements specification. Because of this communication problem, it is claimed that no effective verification procedures can be established, and that misinterpretations of the system requirements specification are not discovered. On the other hand, many formal methods do provide good means for specification animation and validation with respect the users requirements. Animation makes it possible to execute the specifications as early prototypes of the systems to be developed. Furthermore, animation enhances experimentation of the requirements, a property which may be useful in projects where the requirements can not initially be stated in a complete and precise way. It is therefore not really appropriate to discuss whether a formal specification can be easily understood, but rather whether it is communicable.

The use of graphical techniques is a widely accepted approach to making specifications more comprehensible. This is well known in areas like data modelling and specification of information systems. In some very limited application areas, product specification profits from the use specialized diagrams and tables that conform to standard notation in their respective areas. In the context of formal software specification, an important consideration is how graphical descriptions and textual specifications can be combined in a way that benefits the specification process. If the formality or generality of e.g. functional block diagrams

could be improved by establishing a relationship to more general formal specification languages, this appears to be a bridgehead for formal methods technology transfer to industry.

In relation to the research activities at the OECD Halden Reactor Project, many member organisations have expressed an interest in research focusing on the combination of graphical and textual notations in formal software specification. This is in agreement with the guidance by the recommendations from the Halden workshop meeting in 1994 on licensing issues of software in programmable units in nuclear power plants. The workshop meeting considered as particularly important the development of graphically based interfaces for software analysis and design activities, including interfaces to make the use of formal methods more user friendly. On basis of these recommendations, the Halden Project initiated two different research activities. The first of these activities has focused on the combination of graphical and textual notations in formal specification, while the second activity has focused on the development of a graphical front-end to the editing of (textual) algebraic specifications. These two activities are described in the following. The use of graphical specification languages has also been studied in the INF-FS project, see section 3.

2.1 Petri Nets and Algebraic Specifications

The research activities on the combination of graphical and textual notations in formal specification was initially carried out in a co-operative project between ENEA (Italy) and the OECD Halden Reactor Project. The aim of this project was to contribute to a clarification of the relationship between graphical descriptions and formal specifications, and to provide guidelines for how they can be combined in order to utilize the strengths of each approach. One of the project assignments was to investigate how graphical descriptions could be supported by the algebraic specification language and associated tool (the HRP Prover) developed at the Halden Project. Since many graphical description languages can be translated to Petri nets, the focus of the investigations has been put on the translation of these nets into algebraic specification. Petri nets represent in this context an intermediate language between graphical descriptions and algebraic specifications. Furthermore, the importance of Petri nets in the nuclear sector is well documented through applications such as fault diagnosis [22] and fault detection [34] [35] in nuclear reactors, fault tolerance in nuclear reactor protection systems [8], and modelling of work flow in nuclear waste management [29]. A characteristic of Petri nets [33] [36] is that they are at the same time state and action oriented, in the sense that both the states and the actions are explicitly described.

An important feature of the new HRP Prover [39] concerns two related classes of algebraic specifications that capture the concept of state. The *state-based* specifications model the state explicitly, while the *transition-based* specifications model the state implicitly by constructing a traditional generator basis. It can be demonstrated that specifications in each of these classes can be transformed into specifications in the other class, while preserving the proven properties of the specifications. As a practical consequence, the specifier can readily re-organize his specification so that it conforms to the most efficient or familiar approach, or to what appears to provide the best starting point for designing and implementing the specified system. There are good reasons for insisting on the need for both state-based and transition-based algebraic specifications. Many of the differences between a state-based and a corresponding transition-based specification emerge when extending, combining, transforming or refining specifications, or when using tool support. Some of these activities may in practice demand one particular of the two specifications, and consequently a need to transform specifications to suit these needs. The need to transform specifications arise in particular when *combining* a state-based and a transition-based specifications. The style of the combined specification is determined by the desired focus in each concrete case, but the combination will anyhow involve a transformation of one of the specifications.

The concepts of state-based and transition-based specifications have been applied in the establishment of a uniform approach to the translation of a wide variety of *autonomous* and *non-autonomous* Petri nets into algebraic specification. The approach involves translating Petri nets optionally into state-based or transition-based algebraic specifications, and using automatic transformation between these two classes in order to utilize their relative merits. The translation makes it possible to analyse the nets with techniques established for algebraic specification, including the use of the HRP Prover, see Figure 1.

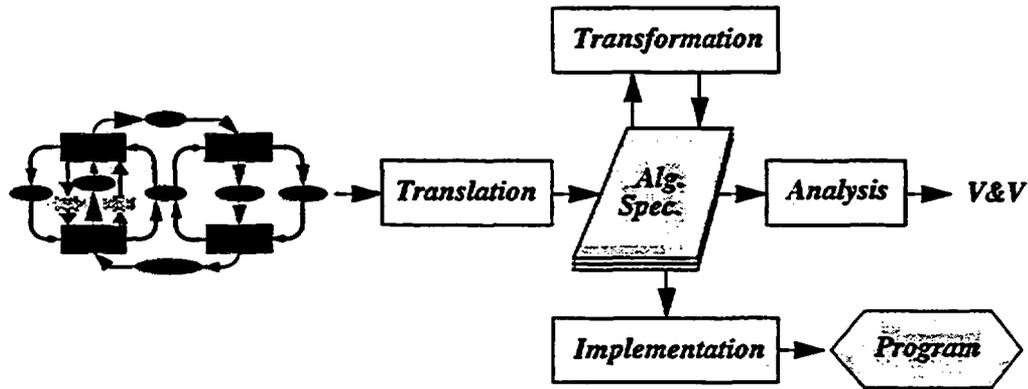


Figure 1. Translating Petri nets into algebraic specifications.

Furthermore, the nets can be manipulated indirectly through transformation of the resulting specifications. In this way, Petri nets and algebraic specifications can be integrated to form a common specification document to be used as a basis for the software development. The approach has been generalised to many different kinds of Petri nets including

- *Autonomous nets*: Condition/Event systems, Place/Transition nets, Coloured Petri nets.
- *Autonomous nets with increased descriptive power*: Inhibitor arc Petri nets, Priority Petri nets.
- *Non-autonomous nets*: Synchronized Petri nets, Timed Petri nets (weak and strong time semantics), Mixed synchronized and timed Petri nets, Interpreted Petri nets.

The choice of state-based or transition-based specification relates to pragmatic concerns about the relative merits of the two classes of algebraic specifications, both in general and with respect to the translation and extension of Petri nets. When it comes to non-autonomous nets, the introduction of synchronization favours the state-based specification, since new external events can then be included successively by introducing new functions. On the other hand, the introduction of timed places favours the transition-based specification, since timing constraints can then be added or modified successively by introducing or modifying separately defined functions. Also the introduction of data processing favours the transition-based specification, since data variables can then be added or modified by introducing or modifying separately defined functions.

Some of the most important findings from the research activities on Petri nets and algebraic specifications can be summarised as follows.

- The algebraic specification resulting from the translation can be used to prove general invariants by means of induction on the transitions.
- The algebraic specification of a Petri net can be gradually extended to incorporate new functioning rules.

- The basic approach to the translation of Petri nets into algebraic specifications can be generalised to a wide variety of Petri nets.
- Automatic transformation between state-based and transition-based specifications gives a flexible and efficient approach to the combination of Petri nets and algebraic specifications.

The work on Petri nets and algebraic specification represents a promising approach to the integration of specifications written in different languages. The approach appears to be applicable to a wide variety of graphical and textual specification languages. Furthermore, the transformation between state-based and transition-based specifications can be utilized in the combination of specifications written in different *textual* languages, such as Z notation [40] and Larch Shared Language [16]. On basis of these observations, it is expected that the approach will prove useful in future work on combination of specification techniques.

2.2 A Graphical Front-End to Algebraic Specifications

The development of a graphical front-end to the editing of algebraic specifications has resulted in a prototype tool called GRAPHIC-AL. Through the use of graphical and hypertext based presentation techniques, GRAPHIC-AL provides useful visualization in the process on building and organising algebraic specifications. The tool can be used as a stand-alone program or as an add-on to the HRP Prover. Concretely, GRAPHIC-AL consists of two main parts:

- The Hypertext Editor: Facilitates navigation in a set of data type specifications; includes a syntax checker.
- The Hierarchy and Version Control Unit: Shows the hierarchical dependencies between data type specifications and allows to work with different versions of a type specifications.

The main goal of the current version of GRAPHIC-AL is to demonstrate an approach to version control and hypertext-based editing of algebraic specifications. In further work, the GRAPHIC-AL will be integrated with a graphical user interface to the HRP Prover.

GRAPHIC-AL works with

- type specifications, i.e. specifications of single (parameterized) types; and
- specifications, consisting of a set of type specifications.

The Hypertext Editor includes several features for creating, opening, saving, checking, and navigating both of these classes of specifications.

The user interface has been developed using Motif widgets (user-interface elements based on the X Windows system), and thereby incorporates a conventional menu bar, scrolled windows, etc. In addition to the menu bar and two editor windows, the user interface provides a hierarchy browser and an information field. While the first editor window provides the means for editing a data type specification, the other editor window is mainly used for information on other data type specifications. The hierarchy browser shows the dependency between all the loaded data type specifications. Finally, the information field gives feedback to the user.

Development of GRAPHIC-AL were carried out by the use of Lex (lexical analysis), Yacc (parsing), and Motif. By way of example, the hypertext widget extends the Motif text widget with right mouse button operations dedicated for the control of GRAPHIC-AL. In general however, writing a new widget by extending an existing one proved to be rather difficult and time consuming, and involved much C programming. As an alternative, Visual Java Development Tools appear promising for the construction of user interfaces and

their integration with the rest of the program.

As a general conclusion, the project GRAPHIC-AL provided a demonstration on how to improve the navigation when editing algebraic specifications. In particular, a rather simple, but effective way was found to present the hierarchical dependencies between data types.

3 Formal Methods and Traditional Software Development Techniques

Recently, the application of formal software development methods have been treated with increasing interest within the nuclear society. The use of formal methods are frequently proposed as a means for developing high-integrity software based systems. The high degree of completeness, dependability, and freedom from defects required for these systems call for effective methods for achieving and demonstrating these quality aspects. Much discussion and, to a certain degree, controversy arose from the verification and validation of the computer-based Darlington shutdown system [10]. Nevertheless, there is today a growing consensus within the nuclear society that more practice on the use of formal methods is needed in order to evaluate their applicability [46]. Several independent studies suggest that there is a need for a systematic, rigorous effort in establishing design requirements to minimize errors in the final product [4]. Licensing authorities in general have a particular interest in representative applications of existing formal methods to make decisions on whether the use of formal methods should be required, which formal methods should be used, what is the appropriate way to use them, and what to require to be formally verified. Much of this motivation comes from the limited value of traditional methods. Following [26], "traditional software-development techniques usually do not provide the levels of dependability demanded by safety-critical systems, and the quality criteria are usually such that the amount of testing that is feasible cannot demonstrate that the desired goals have been achieved". As a matter of fact, there are several important aspects which make the application of software in safety-critical applications fundamentally different from their application in other areas. Safety-critical applications must work when needed, and it is not appropriate to wait for evaluation during use to bring the reliability up to an acceptable level. The realization of the potential benefits of computer-based control and safety systems for nuclear power plants therefore requires *verifying* the reliability of these systems. Traditionally this has been done by means of simulation of the hardware design and exhaustive software testing. It appears however that the use of formal mathematics, in some form, is necessary in order to achieve substantial improvements in the development of dependable software.

Formal methods are useful in demonstrating many principles related to the development of computer-based systems, such as correct and traceable specifications, fault avoidance, structured software development processes, and complete V&V. In general, the provision of a formal specification improves preciseness, because it disciplines the specifier to state explicitly the information necessary to determine what is intended in a particular circumstance. But there are also a broad set of principles for which the use of formal methods need to be combined with other means of demonstration. Examples of such principles are conformance to well-defined standards, competent staff and team organisation, quality assurance, and attention to human factors. The preferred use of formal methods usually also involves support tools, which need to be properly validated for use in the given context. It is regrettable that very few research and development activities have given serious attention to the formal development of these tools. An exception is the activity at the OECD Halden Reactor Project on the formal development of a new version of the HRP Prover [39].

We have already seen how the use of formal methods can be facilitated by the use of graphical notations and techniques. A related approach is to combine formal methods with more traditional software development techniques. This is described in the following subsections, which describes a research activity on the combination of formal methods and CASE tools. Particular emphasis is given to the possibilities of meas-

uring the effect of formal methods in real-life software development.

3.1 Integration of CASE-tools for Formal Methods

The project INT-FS (Integration of formal specification in HAMMLAB 2000) was initiated in 1997 with the aim to experiment with formal methods in the development of HAMMLAB 2000 (see [47] for an introduction to the HAMMLAB 2000 project). Promising formal methods have been identified, and different CASE-tools supporting these methods have been evaluated. It was decided to concentrate on graphical specification languages, like SDL [49], Statecharts [17], MSC [50], and UML [48], that have already proved successful within many companies. In particular, experimental development work has focused on the use of the SDT tool produced by the Swedish company Telelogic. The experience from the application of formal methods and the SDT tool provides important insight into the relationship between semi-formal description techniques and more conventional formal methods. SDT offers a number of facilities, for instance:

- graphical editing of MSC, SDL and UML specifications;
- graphical animation of specifications;
- automatic validation based on exhaustive search;
- complete code-generation towards C.

A first experimental development based on SDT focused on the design of a communication manager for the FAME (Functional Allocation Methods) project. The overall aim of this project has been to develop a framework for analytical studies of important issues in the design of man-machine systems. One such issue is the function allocation between the operator and the process control system. The FAME framework consists of three main parts, an operator component, the communication manager (FCM), and an automation component. FCM is coupled to a simulator of the physical process, and distributes data to the operator module or to the automation module, and back to the process. Further activities in the INT-FS project have concentrated on making a formal specification of the integration platform in HAMMLAB 2000.

Another objective of INT-FS is to give recommendations on how to train personnel in the use of formal methods. The system development within INT-FS is therefore carried out by engineers without former background in formal methods. Because formal methods have a reputation of being hard to comprehend and difficult to apply, system engineers generally tend to be reluctant to using them. The experiences from the project indicate that formal methods of the kind considered are suited as communication media for the different parties involved in a system development. The chosen methods are all graphical, and the CASE-tool SDT is directed towards graphical editing and animation. The developers found that graphical diagrams were well-suited as a medium for discussions with the customers, and that it was easier to restart from a graphical design specification than from conventional code after a longer break. It was also noted that errors were easier to find with animation than with traditional debugging of C code.

3.2 Measuring the Effect of Formal Methods

One important objective of the INT-FS has been to measure the effect of formal methods [43]. For this purpose a procedure for the collection of statistical data (error and work-hour reporting in all development phases) was established. Collected data provides important input to evaluating how the use of formal methods influences on the amount of distribution of errors introduced or detected during the different phases of a development.

Based on the experiences from the experimental development in the INT-FS, several problems were iden-

tified in relation to metrics based on error reports. These problems relate both to the need for an appropriate definition of an error, and to the notion of satisfaction:

- The definition of error must allow for the kind of trial/failure experiments that usually is considered a necessary and desirable part of software development. Moreover, this definition should not depend upon a software process that does not mirror how software is developed in practise.
- Any definition of error is highly dependent on some notion of satisfaction, i.e. what does it mean that a specification or implementation satisfies some requirement imposed by a more abstract specification? This notion of satisfaction must be sufficiently liberal to allow software to be developed in a natural manner; it must be clearly defined, and it must be expressed in such a way that it can be understood and used by the system engineers.

At present, the evidence on experimental work on the effects of formalization is very limited. Two notable exceptions are [5] and [12]. In order to compare conventional and formal developments, care must be taken to ensure that it really is the effects of formalization that is measured, and not the effects of something else. By way of example, the FCM development involved a validation of the formal vs. the informal requirements specification based on a specific review process involving all involved parties. The central question is: To what degree should the activities in the formal development process be mirrored in the conventional one, and vice versa? If the purpose is to measure the effect of formalization only, it appears that the experiments should be organized in such a way that the formal and conventional development differ only with respect to the formalization and the immediate effects of the formalization. This is in agreement with [30], which states that process improvement initiatives should be evaluated by way of comparison between two situations where product and process variables are held constant except for the one whose effect is being checked. Since this may be infeasible across development projects, [30] proposes that such comparisons instead are done across subsystems of the same project.

The FCM development was based on the waterfall process, with the intention to complete each stage before the next was initiated, since this was basically required by the definition of error. In the FCM development, the requirements stage was completed before any of the other stages were started up. For practical reasons however, the design and implementation stages overlapped in time. This does not necessarily mean that the definition of error need to be modified. Rather, the problem appear to relate to the very strict implementation of the waterfall process. The overlap of the development stages occurred because the development of clearly separated sub-components progressed at different speeds. The experience with the FCM development will however be used in the adaption of the error collection routines to an iterative, component based development process in the tradition of [7] and [21].

Other important issues include the definition and selection of adequate metrics, and how to motivate the system engineers to produce the required error reports, and use them in a consistent manner. A challenging task is to set up experiments in such a way that interesting and scientifically valid conclusions can be drawn. Further work of this research activity will therefore focus on improving the techniques and strategies for this kind of experiments, with the aim of providing more significant experimental evidence with respect to the effects of formalization.

4 Theorem Proving and Model-Checking

Formal verification uses mathematical or algorithmic methods to prove the correctness of software specifications with respect to more abstract specifications or desired properties of the software. Much of the motivation behind the use of these methods is their ability to uncover errors, misunderstandings, or unexpected

properties which could easily escape other means of scrutiny. The objectives of the VV-FT project (Verification and Validation using Formal Techniques) is to investigate verification techniques, available verification tools and their applications, in order to propose a framework to support the development and validation of distributed systems. The investigations have concentrated on theorem proving and model checking, which are the two major approaches to formal verification.

Basic to the formal verification of software specifications is the viewpoint that they specify possible *runs* of the software, where each run can be represented as an infinite sequence of states. Accordingly, a property of the specification is understood as a set of runs. With this viewpoint, we can distinguish two classes of properties:

- *Safety properties* express that undesired actions will not happen. A run satisfies a given safety property if every initial (finite) sequence of the run can be extended to a run satisfying that property. As a special case, an invariant property is a property that holds in every state of the execution.
- *Liveness properties* express that desired actions will eventually take place. A given property is a liveness property if it is possible to extend every finite sequence to a run satisfying that property.

Techniques for proving safety properties are typically based on some form of generalized mathematical induction.

Model-checking as a technique relies on building a finite model of a system and checking that a desired property holds in that model (using some variant of reachability analysis). In contrast, theorem proving is based on expressing both the system and its desired properties as formulas in some mathematical logic. The logic is given by a formal system, which defines a set of axioms and a set of inference rules. Theorem proving is the process of finding a proof of a property from the axioms of the system. The suitability of each of the two classes of techniques depends on the characteristics of the analysed system:

- Model-checking techniques are applicable to systems whose states have short and easily manipulated descriptions. Typical systems in this category are those whose intricacy resides more in the control than in data. These are systems whose role is more readily described by their possible interaction sequences than by the transformations they apply to complex data.
- Theorem proving techniques are more general than model-checking, and can be used to verify many types of systems including those whose role is described by the transformations they apply to complex data.

Due to their respective advantages and disadvantages, a combination of the techniques may be needed for large complex systems. By way of example, interactive theorem proving tools may help system developers to verify decomposition and abstraction steps, while model-checking tools may be used to handle relatively small and decidable subsystems.

In the VV-FT project, several tools within each of these classes were compared. The theorem proving tools included in this study were the ACL2 system [25], the HOL system [15], the HRP Prover [39], the LP Prover [16], and the PVS system [11]. The model-checking tools were the CWB-NC system [9], the SMV system [28], the SPIN system [18], and the UV system [24]. The evaluation criteria put special emphasis on the applicability of the tools for the development of distributed software. As interactive theorem proving tools were considered, the PVS system was considered to be better than the other systems for the specification of distributed software systems. Due to its flexibility, the HOL system might be a better choice for proving theorems. Another advantage with the latter tool is that it appears to be easier to extend with derived proof rules. As model checking tools were considered, the SPIN system was considered to be better than the other

tools for specification and verification of distributed systems.

The further activity in this project has been to investigate whether it is possible to apply this methodology on the verification of operator procedures. In a preliminary study, SPIN was applied on a simple example procedure. Based on this experience, a case study was initiated using this tool on a real procedure. The purpose of this study was to find out how well SPIN handles relatively large procedures, and what kind of properties can be verified. The activities include modelling the procedure in Promela, which is the language of SPIN, modelling the environment to which the procedure is supposed to be applied, and investigating the types of errors that could possibly be detected by using SPIN. Based on related work at the OECD Halden Reactor Project, further work will also look into the possibilities of combining this approach with qualitative simulation [27]. The latter approach puts more emphasis on the modelling of the qualitative behaviour of the underlying plant, and a combination therefore appears promising in view of the extended scope of verification.

The combination of model-checking and theorem proving is a relevant issue also in the research activities on Petri nets and algebraic specifications (see section 2), where inductive theorem proving complements traditional net analysis techniques. One of the main advantages of the induction approach is its generality and the flexibility it allows for the formulation of invariants. This is first of all due to the fact that the inductive approach easily applies to all specifications classified as *state-based* or *transition-based* algebraic specifications [39].

5 Program Testing Techniques

Several demonstration principles for software based systems important to safety refer to the importance of testing in relation to operational testability, the required system functions, etc. In general, to test a program is to execute it with selected test data to demonstrate that it performs its task correctly. Testing is an essential part in the assessment of a software product, and complementary to other V&V activities. Ideally the test data should be selected so that all potentially residual faults are revealed. However, exhaustive testing is in general not possible, so the optimal test strategy is the one which maximises the probability of revealing all possible residual program faults. In some cases, testing may be facilitated by the use of *equivalence partitioning*. This is a technique for determining which classes of input data receive equivalent treatment by a system, a software module, or program. A result is the identification of a finite set of software functions and of their associated input and output domains [13].

Effort-intensive testing activities can be viewed as a necessary complement to the more process-oriented development procedures, since these procedures are regularly considered incapable of providing the necessary confidence in the program. Testing has the advantage of giving reliability figures directly, and provide flexibility in the choice of the number and distribution of test data. A basic problem related to testing is to determine when to *stop* the testing. Several objective criteria have been proposed, based on the test sets, the number of bugs detected, the marginal detection rate, or the target reliability level (see [44]). Testing also requires a method to decide whether the result of a computation is correct or not, i.e. an *oracle*. A frequent problem with testing practise is a poor consideration of the quality of the process and techniques employed during the software development. The intrinsic discrete nature of computer hardware and software also makes testing difficult and sometimes unreliable, because the untested sequences of states do not necessarily lie in the "neighbourhood" of tested sequences. If the system is non-deterministic, the system will not even necessarily repeat its observed behaviour on a test case.

5.1 Software Sensitivity Analysis and Reliability Assessment

The objective of the EISTRAM (Experimental Investigation of Software Testing and Reliability Analysis Methods) project has been to investigate test based measures of software dependability [14]. One such measure is the sensitivity of a program, i.e. the probability that a program fault will lead to a failure during execution. In the EISTRAM project, the PIE technique [45] has been investigated as one method for sensitivity analysis. It combines execution-, infection- and propagation analysis to provide an estimate on where possible residual faults may hide. Combined with software reliability measures, such as time to failure, it should be possible to increase the confidence in the fault freeness of the program in cases where no failures have been revealed during testing. The technique has been applied on programs developed in other research activities at the Halden Project, including the PRM program developed with the use of formal methods in the EvalFM project [37], and a program that was developed in the Project on Diverse Software (PODS) [6].

The acronym PIE stands for Propagation, Infection and Execution, which during the analysis are performed in reverse order, i.e. execution of a location, infection of the data state, and propagation of a fault to a discernible output. While the purpose of most mutation testing techniques is to prove the absence of certain classes of faults, correctness is not an issue with the PIE technique. Instead the purpose is to identify locations in a program, where faults, if they exist, are more likely to remain undetected during testing. The locations of interest are primarily assignment statements, input/output statements, or the condition part of an IF, CASE, UNTIL, or WHILE statement. The technique is closely related to three conditions that are both necessary and sufficient to cause a software failure, and that must occur in the following sequence:

1. An input must cause the fault to be executed, i.e. the faulty location must be reached and executed.
2. Once the fault is executed, the succeeding data state must contain a data state error, i.e. the succeeding data state must contain an incorrect variable/value pairing.
3. Once the data state error is created, it must propagate to an erroneous output state, i.e. the data state error must cause an incorrect output from the program.

For each investigated location, the PIE analysis will produce one set of probability estimates for each of these conditions. The PIE technique then uses the minimum estimates from each set to estimate the sensitivity for each location. This is an estimate of the probability that a fault will cause a failure at a particular location under a specified input distribution:

- If high sensitivity is observed, there will be a high probability that faults, if they exist, will be revealed during testing.
- If low sensitivity is observed, it is likely that faults will remain undetected during random testing, and the location will be a candidate for further investigation.

Through a limited literature survey, a fault analysis of previous projects, and accumulated experience during the experiments, the EISTRAM project classified the mutants into 23 main operators. Specific interest was given to mutants and mutant operators that caused the zero infection estimates. As a part of the EISTRAM project, programs were developed to automatically generate most of the mutants based on these operators. The list of simulated faults were narrowed in order to reduce the time required for performing the testing, and because it was considered desirable to only simulate those faults that are likely to be made by a programmer. For this purpose, ten fault criteria were defined.

The test case based on the PRM program consisted of 122 locations from the subroutine part of a Pascal program for performing a power range monitoring of a nuclear reactor. In the analysis of the program, the PIE technique was applied to the 122 locations by initially testing with a normal input distribution. It was

observed that 66 locations had zero sensitivity estimates. 55 of these had a zero infection estimate, where 849 out of 5801 tested mutants caused a zero infection estimate. By applying the fault criteria to the 849 mutants, it was observed that the number of locations with a zero infection estimate was reduced to 29, and the number of mutants were reduced to 87. Only one location now had both a zero infection and zero propagation estimate. The observations from the analysis of the PRM program were in agreement with the observations from smaller test cases. It was also observed from the list of locations with zero infection estimates that only 7 of the 44 analysed subroutines had a large proportion of such locations. Four of these subroutines were functions of type Boolean, and three were procedures that returned used-defined records, of which several fields were of type Boolean.

By applying the PIE technique to the various test cases, it was observed that the number of locations which are likely to hide a possible fault during random testing is very high. The number of locations was reduced by using several input distributions to test the mutants. However, the number of locations were still high, approximately half of the tested locations. The use of mutant selection criteria can reduce the number of mutants that give zero infection estimates, and thereby the number of locations likely to hide faults. The inherent danger is that the application of these criteria may result in the removal of mutants that represent likely program faults. The high number of locations that would be likely to hide a fault during testing also means that one has a large number of pinpointed locations which are candidates for other testing methods. In this view one could conclude that using the PIE technique was not very efficient. On the other hand, the large number of "insensitive" locations could be an indication of a fault tolerant program.

6 Software Process- and Product Quality

Generally speaking, there are two principal ways of ensuring software quality - one in terms of process, and another in terms of product [38]. In a process-oriented approach, quality is seen as an outcome of a good software development process. In a product-oriented approach, quality is assessed or ensured by directly evaluating a given piece of software. In both cases, the aim of the quality assurance activities is to ensure that the software exhibits properties like correctness, reliability, safety, efficiency, and maintainability. If we chose to use a process-oriented approach to ensure e.g. correctness, we are in reality making the assumption that the quality of the process will at least help in ensuring this aspect of product quality. We may still employ various testing strategies as part of the process, but these are then typically required in the process in terms of testing plans, acceptance criteria, etc. Alternatively, we could use a product-oriented approach, and focus more directly on analysing the software being developed. This could be done experimentally in terms of program testing, or more analytically through formal verification.

Intuitively, it is easy to see that neither the process-oriented nor the product-oriented approach is fully satisfactory. By way of illustration, the *process-oriented approach* may fail due to over-emphasis on traditional quality assurance activities on the cost of adequate testing and verification. On the other hand, the *product-oriented approach* may fail due to insufficient consideration of how the testing and verification can be facilitated by controlling the software development process. These pitfalls immediately suggest that an optimal approach requires a combination of the process-oriented and the product-oriented approach:

- The process aspect should be maintained by following a development process that can be planned, tracked, and reviewed.
- The process aspect should be complemented by guidelines to "design for V&V", i.e. the development process should promote the development of programs that are structured in a way that facilitates testing and verification of the final products.
- The review/audit activities should cover the question of adequate V&V.

Measuring Quality

In order to facilitate specification and measurement of software quality, several attempts have been made to identify what characterizes a high quality software product. This is of vital importance both in a product-oriented and in a process-oriented approach to software QA. While product-metrics tend to focus on single modules or single quality characteristics, process-metrics usually concern the project itself or specific phases in the development. Unfortunately, it does not seem likely that a standard set of metrics can be established. Neither is there any general consensus about what are the best measurement practices. There are however a few paradigms that appear to have great impact on the development of international standards for software development. Leading paradigms are Goal Question Metrics [3], Quality Function Deployment [2], and Software Quality Metrics [19].

Process-Oriented Quality Assurance

As was stated above, the process-oriented approach to software quality is based on viewing quality as an outcome of a good development practise. Some of the leading process-oriented quality assurance frameworks are Total Quality Management, the Capability Maturity Model [32], and Risk Management [7]. A general weakness of process-oriented approaches to software quality is insufficient evidence on their actual influence on product quality. By way of example, it remains to be demonstrated that companies with high ratings with respect to CMM in fact produce software of higher quality. It appears that a strong focus on conformance to a model or standard tends to underrate the importance of adequate V&V. This is also the case with quality systems based on standards proposed for use in software quality assurance. The vast majority of these standards reflect the process-oriented approach to software quality assurance. Due to a strong focus on certification, there is a clear tendency to put too much emphasis on ISO 9000 and 9000-3, while ignoring other standards. According to [42], no empirical evidence, no theory, and no explicit model has been given that justifies or explains the relation between the ISO 9000 family and the accomplishment of improved software product quality. In fact, a large empirical survey among European software suppliers shows that only very few managers are able to quantify the benefits of an ISO 9000 quality system, including its impact on product quality [41].

Product-Oriented Quality Assurance

In the product-oriented approach to software quality, various quality aspects are assessed or ensured by directly evaluating given pieces of software. By way of example, software development frequently involves exhaustive testing of specifications and code. It is possible to view these effort-intensive activities as a necessary complement to the more process-oriented development procedures, since these are regularly considered incapable of providing the necessary confidence in the program. Testing has the advantage of giving reliability figures directly, and provide flexibility in the choice of the number and distribution of test data. A basic problem related to testing is to determine when to stop the testing. Testing also requires a method to decide whether the result of a computation is correct or not, i.e. an *oracle*. In relation to process/product quality, a frequent problem with testing practise is a poor consideration of the quality of the process and techniques employed during the software development. Finally, the intrinsic discrete nature of computer hardware and software makes testing difficult and sometimes unreliable, because the untested sequences of states does not necessarily lie in the "neighbourhood" of tested sequences. If the system is non-deterministic, the system will not even necessarily repeat its observed behaviour on a test case. Other product-oriented quality assurance techniques include software reliability prediction, program verification, fault tolerance, and diversity. All of these techniques have their strengths and weaknesses with respect to their effectiveness in ensuring product quality. Nevertheless, their actual use typically fall short of capturing the impact of process-oriented quality assurance techniques.

6.1 Synthesized Quality Assessment

For a final acceptance of a safety-critical software-based system, a thorough safety assessment is necessary. For each application in nuclear power, this is represented by a safety case that must be put forward for approval by the licensing authorities. The safety case aims at demonstrating that the requirements specification is correct and complete, and in all aspects satisfied by the delivered system. But the safety case must also cover other aspects related to the system, such as the development process and its conformance to relevant standards, and required performance of the system throughout its operational life [1]. The basis of a safety case consists of many different types of evidences, some of which are of a qualitative nature. Following [13], the following three kinds of evidences need to be produced:

- Evidence related to the quality of the development process.
- Evidence related to the adequacy of the product.
- Evidence of the competence and the qualifications of the staff involved in all the phases of the system life cycle.

Since not all aspects that influence the confidence one can have in a program are measurable (in statistical terms), the safety assessment calls for some approach to combining evidences of a disparate nature. This is especially the case with systems based on pre-existing software, including commercial, off-the-shelf software (COTS). As is stated by [13], there are several issues involved with use and validation of pre-existing software:

- The functional and non-functional behaviour is often not clearly defined and documented.
- The documentation and the data on operational experience are often not adequate enough to provide the evidence which would be required to compensate for the lack of knowledge on the product and on its development process.
- As a consequence of the two previous issues, acceptance criteria and procedures of investigation for demonstrating fitness for purpose for a specific application may be difficult to put in place.
- The operational experience may not be in exact correspondence with that of the intended application. Therefore, software paths of unknown quality may be invoked by the application.

Aspects like producer's pedigree, the software process being employed, etc., need to be considered, although in a different way than evidences like code complexity, test results, etc. Many of these evidences relate directly to process- or product-oriented quality practices and techniques. A systematic approach to measuring and combining influences would therefore also represent an integration of process/product quality in safety assessment. Furthermore, the same framework would provide useful guidance in the development of the system. In particular, the use of sensitivity analysis could be used to find the relative importance of the different types of information. If software product quality was to be considered in isolation, we would need detailed information about the software. Since this is not always available, we may have to focus on evidences giving some indication on these aspects. By way of example, conclusions on the level of code complexity or well-structuredness of a program may to some extent be drawn from information on functionality, modes of operation, etc. There is also a great potential benefit in utilizing knowledge about how specific development processes and techniques will influence on such factors.

A qualitative type of reliability measure is expressed as a subjective judgement, as a "belief" in fault freeness. A methodology which has been proposed is to use Bayesian Belief Nets (BBNs) and engineering judgement to combine evidences from different information sources for a qualitative assessment of this belief. The objective of using BBNs in software safety assessment is to show the link between basic informa-

tion and the confidence one can have in a system. One method for using BBNs in predicting software quality is described in [31]. The BBN methodology is not only applicable in the final assessment of a product, but can also be used to show the achievement of subgoals throughout the whole software life-cycle. The method should therefore be applicable for different purposes, as e.g. to evaluate COTS systems, assess the development process until a final assessment of the system, and even assess the operation and maintenance of the system. A simple illustration of the methodology can be given in terms of Figure 2, where e.g. reliability can be represented by a *target node*.

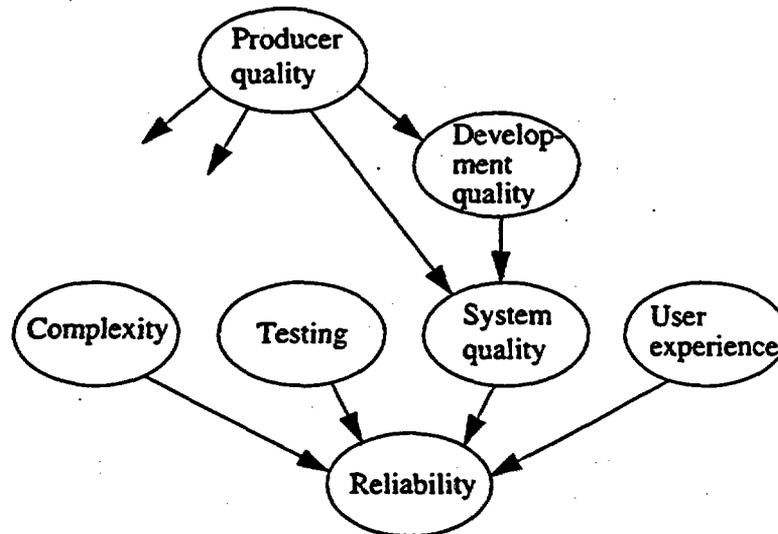


Figure 2. Illustration of the BBN methodology.

In an extended network, reliability would typically influence safety assessment, represented by yet another target node. Examples of *observable nodes* are complexity of code and testing, for which exact numbers may be available. But there are also other factors influencing reliability, such as system quality which again is influenced by development quality. One obvious example would be the use of formal methods. We may have only limited information on these factors, and the nodes are therefore classified as *intermediate nodes*. For the associated variables, we will then give *a priori* values represented by discrete numbers or distributions. Based on the assignments given to each node and edge in the network, we can use a computer program to compute a "belief" value for each target node variable.

In spite of a large amount of published papers on BBNs and related graphical probability models, it is difficult to find evidences on the application of BBNs for real or realistic safety cases. In order to achieve experience on such use, the OECD Halden Reactor Project will in co-operation with member organisations apply the proposed methodology on a real or realistic system. The project assignments selected to achieve this objective first of all rest on the selection of a programmable protection system based on a configurable software system. This protection system will be used as a test case in the project. Particular criteria on this system will then be deduced from the plant wide safety criteria, followed by a collection of all available information relevant for the evaluation of the software. By way of a Bayesian net, all information sources will be connected to a target node directly relevant for approval of the system, where different methods will be used to assign numbers to the nodes and edges of the net. Finally, a commercially available tool will be used to make computations of the net. Particular emphasis will be given to sensitivity analysis with respect to the assigned numbers, with the aim of finding the relative importance of the different types of informa-

tion.

In order to make optimal use of evidences, these will often have to be detailed further. In this way, more aspects can be made visible and thereby facilitate quantitative assessments. This can be illustrated by the evidence represented by the use of formal methods. In order to assess the influence of such use on the quality of the software product, several factors need to be considered, such as

- Which methods have been used?
- How have these methods been used?
- How well is the use of these methods documented in this particular project?

Visibility is further improved by assessing the influence on quality in terms of the various quality characteristics, such as those defined in ISO 9126 [20]. In this way, more detailed measures are identified in a systematic way, some of which may be appropriate for quantitative assessment. As a consequence, an evidence that is given a subjective assessment in the overall, can possibly be detailed to a combination of subjective and quantitative assessments, and thereby better utilize the available evidence. This demonstrates the need for empirical results on the influence of formal methods on software product quality. It also indicates how these results can be utilized by combining more detailed subjective measures and quantitative measures into an overall assessment on the influence of the use of formal methods for each particular development project.

7 Conclusions

The effectiveness of software quality assurance largely depends on the success of combining complementary notations and techniques. While this is a frequent observation both in research and practise, there is still a lack of consensus on the usefulness of the various combinations. In many instances, even the effectiveness of a single technique is difficult to measure. The Halden Project addresses these problems by investigating how different notations and techniques can be combined in order to improve the overall scope and effectiveness of specification and verification. On basis of these activities, the present paper has discussed some of the options available for combining software quality assurance techniques. While the various options reflect a wide variety of techniques, several classes of combinations involve the use of formal methods. In particular, the combination of graphical and textual notations represents a promising approach to making formal specifications comprehensible to a wider group of users.

The paper has presented research results related to the combination of Petri nets and algebraic specifications, as well as to the development of a graphical front-end to the editing of algebraic specifications. The concepts of *state-based* and *transition-based* algebraic specifications have been applied in the establishment of a uniform approach to the translation of a wide variety of *autonomous* and *non-autonomous* Petri nets into algebraic specification. The approach involves translating Petri nets optionally into state-based or transition-based algebraic specifications, and using automatic transformation between these two classes in order to utilize their relative merits. The translation makes it possible to analyse the nets with techniques established for algebraic specification, including the use of the HRP Prover. Furthermore, the nets can be manipulated indirectly through transformation of the resulting specifications. In this way, Petri nets and algebraic specifications can be integrated to form a common specification document to be used as a basis for the software development. The development of a graphical front-end to the editing of algebraic specifications has resulted in a prototype tool called GRAPHIC-AL. Through the use of graphical and hypertext based presentation techniques, GRAPHIC-AL provides useful visualization in the process on building and organising algebraic specifications. The main goal of the current version of GRAPHIC-AL is to demonstrate an approach to version control and hypertext-based editing of algebraic specifications. In further work, the

GRAPHIC-AL will be integrated with a graphical user interface to the HRP Prover.

The practicality of formal methods is further enhanced by a proper combination of formal methods with CASE-tools and traditional development techniques. The paper has presented results from a research project concerned with using, and measuring the effect of, formal methods in real-life software development. The experiences from the INT-FS project indicate that formal methods of the kind considered are suited as communication media for the different parties involved in a system development. The chosen methods are all graphical, and the CASE-tool SDT is directed towards graphical editing and animation. The developers found that graphical diagrams were well-suited as a medium for discussions with the customers, and that it was easier to restart from a graphical design specification than from conventional code after a longer break. It was also noted that errors were easier to find with animation than with traditional debugging of C code. When it comes to measuring the effect of formal methods, the INT-FS project established a procedure for the collection of statistical data (error and work-hour reporting in all development phases). Collected data provides important input to evaluating how the use of formal methods influences on the amount of distribution of errors introduced or detected during the different phases of a development. Experiences from the INT-FS project do however indicate a need for more sophisticated and reliable error counting techniques that do not conflict with how software systems are developed in practise.

A third class of combinations concerns the use of complementary formal verification techniques. A distinction is made between theorem proving and model-checking, both of which represent well established approaches to the verification of software specifications. The suitability of each of the two classes of techniques depends on the characteristics of the analysed system: Model-checking techniques are applicable to systems whose states have short and easily manipulated descriptions. Typical systems in this category are those whose intricacy resides more in the control than in data. These are systems whose role is more readily described by their possible interaction sequences than by the transformations they apply to complex data. Theorem proving techniques are more general than model-checking, and can be used to verify many types of systems including those whose role is described by the transformations they apply to complex data. Due to their respective advantages and disadvantages, a combination of the techniques may be needed for large complex systems. By way of example, interactive theorem proving tools may help system developers to verify decomposition and abstraction steps, while model-checking tools may be used to handle relatively small and decidable subsystems.

There are also several options available for the combination of program testing techniques. While formal verification is concerned about proving correctness of a specification, testing normally focuses on the (possibly symbolic) execution of a specification or a program. The paper has discussed how testing can be made more focused if it is guided by the PIE-technique, which is a dynamic failure-based technique for performing program sensitivity and testability analysis. The main observation from the experimental use of the PIE-technique was that the number of locations which are likely to hide a possible fault during random testing was very high, also after reducing the number of such locations by applying different mutant selection criteria. The high number of locations also means that one has a large number of pinpointed locations which are candidates for other testing methods or techniques. In this view one could conclude that the PIE-technique was not very efficient. On the other hand, the high number of insensitive locations could also be an indication of fault tolerant programs.

Finally, the paper has discussed the problem of combining the process-oriented and product-oriented approach to software quality. Due to their inherent limitations, neither of the two classical approaches appears to be fully satisfactory in ensuring adequate software quality. This is particularly the case for the development and assessment of safety critical systems. By way of example, the *process-oriented approach* may fail

due to over-emphasis on traditional quality assurance activities on the cost of adequate testing and verification. On the other hand, the *product-oriented approach* may fail due to insufficient consideration of how the testing and verification can be facilitated by controlling the software development process. It seems clear that the quality of a software product should not be seen in isolation from the development process. This is motivated both from the need to ensure an efficient process and from the underlying assumption that quality of the development process has an impact on the product quality. The paper has demonstrated how the integration of software process- and product quality relates to the need to incorporate a variety of factors into quality assurance or assessment. It is demonstrated how these factors, which represent disparate evidences of software quality, can be combined in quality assessment by means of so-called Bayesian networks.

8 References

- [1] AECB, DSIN/IPSN, NII, USNRC. *Four party regulatory consensus report on the safety case for computer-based systems in nuclear power plants*. Health & Safety Executive, Nov. 1997.
- [2] Y. Akao (ed.). *Quality Function Deployment: Integrating Customer Requirements Into Product Design*. Productivity Press, Cambridge, Mass., 1990.
- [3] V.R. Basili. Applying the Goal/Question/Metric paradigm in the experience factory. In N. Fenton, R. Whitty and Y. Iizuka (eds.). *Software Quality Assurance and Measurement: A Worldwide Perspective*, pages 23-44. International Thomson Computer Press, 1995.
- [4] L. Beltracchi. NRC research activities. In D.R. Wallace, B.B. Cuthill, L.M. Ippolito and L. Beltracchi (eds.). *Proc. Digital Systems Reliability and Nuclear Safety Workshop*, Sep. 13-14, 1993. NUREG/CP-0136, United States Nuclear Regulatory Commission, Washington DC, 1994.
- [5] J. Bicarregui, J. Dick, and E. Woods. Quantitative analysis of an application of formal methods. In *Proc. FME'96: Industrial Benefit and Advances in Formal Methods*, Oxford, LNCS 1051, pages 60-73. Springer-Verlag, 1996.
- [6] P. Bishop, D. Esp, M. Barnes, P. Humphreys, G. Dahll, and J. Lahti. PODS - the project on diverse software. *IEEE Trans. Software Engineering*, SE-12(9): 929-940, 1986.
- [7] B.W. Boehm. *Software Risk Management*. IEEE Computer Society Press, Los Alamitos, Calif., 1989.
- [8] G.H. Chisholm, B.T. Smith, and A.S. Wojcik. Formal specifications for safety grade systems. In *Proc. Methodologies, Tools, and Standards for Cost-Effective, Reliable Software Verification and Validation*. EPRI, CA, USA, Jan. 1992.
- [9] R. Cleaveland, J. Parrow, and B. Steffen. The Concurrency Workbench: A semantics-based verification tool for the verification of concurrent systems. *ACM Trans. Programming Languages and Systems*, 5(1):36-72, Jan. 1993.
- [10] R.H. Crane. Experience gained in the production of licensable safety-critical software for Darlington NGS. In *Proc. Methodologies, Tools, and Standards for Cost-effective, Reliable Software Verification and Validation*. EPRI, Palo Alto, CA, USA, Jan. 1992.
- [11] J. Crow, S. Owre, J. Rushby, N. Shankar, and M. Srivas. A tutorial introduction to PVS. In *Proc. WIFT'95: Workshop on Industrial-Strength Formal Specification Techniques*, Boca Raton, Florida, April 1995.

- [12] J. Draper, H. Treharne, T. Boyce, B. Ormsby. Evaluating the B-tool on an avionics example. In *Proc. DAISA '96*, pages 89-97. European Space Agency, 1996.
- [13] European Commission. *European nuclear regulators' current requirements and practices for the licensing of safety critical software for nuclear reactors*. Draft report, revision 8, Luxembourg: Office for Official Publications of the European Communities, 1998.
- [14] B.A. Gran and H. Thunem. EISTRAM - experimental investigation of the PIE-technique. *Proc. ESREL '98*. Trondheim, Norway, June 1998.
- [15] M.J.C. Gordon and T.F. Melham (eds.). *Introduction to HOL: A Theorem Proving Environment for Higher Order Logic*. Cambridge University Press, 1993.
- [16] J.V. Guttag and J.J. Horning. *Larch: Languages and Tools for Formal Specification*. Springer-Verlag, 1993.
- [17] D. Harel. STATECHARTS: A visual formalism for complex systems. *Science of Computer Programming*, 8:231-274, 1987.
- [18] G.J. Holzmann. The model checker Spin. *IEEE Trans. Software Engineering*, 23(5): 279-295. May 1997.
- [19] IEEE Standard 1061. *Software Quality Metrics Methodology*. Available from IEEE Computer Society Press, Los Alamitos, Calif., 1992.
- [20] ISO 9126. *International Standards Organisation. Information technology - Software product evaluation - Quality characteristics and guidelines for their use*. ISO/IEC IS 9126.
- [21] I. Jacobson, M. Christerson, P. Jonsson, and G. Øvergaard. *Object-oriented Software Engineering - A Case Driven Approach*. Addison-Wesley, 1992.
- [22] N.A. Jaleel and H. Nicholson. *Petri Nets and Fault Diagnosis in Nuclear Reactors*. Research report no. 413, Department of Control Engineering, University of Sheffield, Nov. 1990.
- [23] K. Jensen. An Introduction to the Theoretical Aspects of Coloured Petri Nets. In J.W. de Bakker, W.-P. de Roever, and G. Rozenberg (eds.). *A Decade of Concurrency*, LNCS 803, pages 230-272. Springer-Verlag, 1994.
- [24] M. Kaltenbach. *Model checking for UNITY*. Technical Report CS-TR-94-31, University of Texas, Austin, Dec. 1994.
- [25] M. Kaufmann and J.S. Moore. An industrial strength theorem prover for a logic based on Common Lisp. *IEEE Trans. Software Engineering*, 23(4): 203-213, April 1997.
- [26] J. Knight and B. Littlewood. Critical task of writing dependable software. *IEEE Software*, pages 16-20, Jan. 1994.
- [27] B. Kuipers. *Qualitative Reasoning - Modeling and Simulation with Incomplete Knowledge*. MIT Press, 1994.
- [28] K.L. McMillan. *Symbolic Model Checking: An Approach to the State Explosion Problem*. Technical Report CMU-CS-92-131, Carnegie-Mellon University, May 1992.
- [29] K.H. Mortensen and V. Pinci. Modelling the work flow of a nuclear waste management program. In R. Valette (ed.). *Application and Theory in Petri Nets 1994. Proc. 15th Int'l Petri Net Conference*, LNCS 815, pages 376-395. Springer-Verlag, 1994.

- [30] J.D. Musa and W.W. Everett. Software-reliability engineering: Technology for the 1990s. *IEEE Software*, pages 36-43, Nov. 1990.
- [31] M. Neil and N. Fenton. Predicting software quality using Bayesian belief networks. In *Proc. 21st Annual Software Engineering Workshop*. NASA/Goddard Space Flight Centre, 1996.
- [32] M.C. Paulk, B. Curtis, M.B. Chrissis, and C.V. Weber. *Capability Maturity Model for Software, Version 1.1*. SEI Technical Report SEI-CMU-93-TR-24, Software Engineering Institute, Pittsburg, Pa., 1993.
- [33] J.L. Peterson. *Petri Net Theory and the Modeling of Systems*. Prentice-Hall, 1981.
- [34] J. Prock. A 3 level conception for signal validation and early fault detection in closed systems. Enlarged Halden Programme Group Meeting (OECD Halden Reactor Project), Bolkesjø, Norway, 1990.
- [35] J. Prock. A new technique for fault detection using Petri Nets. *Automatica*, 27: 239-245, 1991.
- [36] W. Reisig, *Petri Nets, An Introduction*. Springer-Verlag, 1985.
- [37] T. Sivertsen. A case study on the formal development of a reactor safety system. In *Proc. FME'96: Industrial Benefit and Advances in Formal Methods*, Oxford, LNCS 1051, pages 18-38. Springer-Verlag, 1996.
- [38] T. Sivertsen. Integration of software process/product quality. In *Proc. IAEA Specialists' Meeting on Licensing of Backfitting/Modernization of Instrumentation and Control systems Important to Safety*. Vienna, Austria, September 7-9, 1998.
- [39] T. Sivertsen. Putting principles into practise: The formal development of a theorem prover. In *Proc. 26th. Water Reactor Safety Information Meeting*, October 26-28, 1998, Bethesda, Maryland. United States Nuclear Regulatory Commission, Washington DC, 1998.
- [40] J.M. Spivey. *The Z Notation - A Reference Manual*. Prentice-Hall, 1989.
- [41] D. Stelzer, W. Mellis, and G. Herzworm. Software process improvement via ISO 9000? Results of two surveys among European software houses. In *Proc. Twenty-Ninth Annual Hawaii International Conference on System Sciences*, Volume 1, pages 703-712. IEEE Computer Society Press, Washington, 1996.
- [42] D. Stelzer, W. Mellis, and G. Herzworm. A critical look at ISO 9000 for software quality management. *Software Quality Journal*, 6:65-79, 1997.
- [43] K. Stølen and P. Mohn. Measuring the effect of formalization. In *Proc. SAM98, 1st Workshop of the SDL Forum Society on SDL and MSC*, June 29 - July 1, 1998, Berlin, Germany.
- [44] A. Villemeur. *Reliability, Availability, Maintainability and Safety Assessment. Volume 2: Assessment, Hardware, Software and Human Factors*. Wiley, 1992.
- [45] J.M. Voas. PIE: A dynamic failure-based technique. *IEEE Trans. Software Engineering*, SE-18(8): 717-727. 1992.
- [46] D.R. Wallace, B.B. Cuthill, L.M. Ippolito and L. Beltracchi (eds.). *Proc. Digital Systems Reliability and Nuclear Safety Workshop*, Sep. 13-14, 1993. NUREG/CP-0136, United States Nuclear Regulatory Commission, Washington DC, 1994.

- [47] F. Øwre. Achievements and further plans for the OECD Halden Reactor Project man-machine systems programme. In *Proc. 26th. Water Reactor Safety Information Meeting*, October 26-28, 1998, Bethesda, Maryland. United States Nuclear Regulatory Commission, Washington DC, 1998.
- [48] UML proposal to the object management group, version 1.1, September 1, 1997.
- [49] Recommendation Z.100 - CCITT specification and description language (SDL). ITU, 1993.
- [50] Recommendation Z.120 - Message Sequence Chart (MSC). ITU, 1996.

Human Factors Engineering and Control Room Design using a Virtual Reality Based Tool for Design and Testing

Michael Louka, Conny Holmstrøm and Fridtjov Øvre

OECD Halden Reactor Project
N-1751 Halden, Norway
Tel: +47 69 21 22 00 Fax: +47 69 21 22 01
e-mail: Michael.Louka@hrp.no

ABSTRACT

This paper describes a user-centred approach to control room design for the nuclear industry. The establishment of a "Control Room Philosophy" and the use of virtual reality (VR) technology in the design process are key features of this approach. The control room philosophy identifies the functional aspects of a control centre, defining principles and guidelines to be used throughout the design process. It can be viewed as a functional requirements specification that is used to guide the design and development of a control centre. VR technology is used to visualise a design based on a control room philosophy. VR technology is not only used to visualise the design, enabling designers to interactively modify it, but also to test and evaluate it against regulative standards for nuclear control rooms. A design documentation system (DDS) has been integrated with the VR tools to support the documentation of the design process in a structured manner. The use of VR to visualise a control room has enabled control room operators to actively participate in the design of new control rooms together with human factors experts, and a VR of a design can replace a physical mock-up for design testing and evaluation.

1. INTRODUCTION

There is increasing recognition of the need to apply human factors principles to a design at as early a stage in the design process as possible. Virtual Reality (VR) is an interactive medium that assists in achieving this objective because it enables design engineers to examine evolving designs from the viewpoint of the end-user, reducing the risk of costly design-induced operational problems while at the same time improving the quality of the final design.

In the context of the approach described in this paper, VR is defined as an artificial world (the evolving control room design) in which the user (the design engineer and others participating in the design process) can navigate and interact with objects in the world in real-time. The virtual control room is rendered using three-dimensional computer graphics.

A control room philosophy provides the foundation for developing a control centre concept based on functional and operational requirements. It identifies the roles of operators in addition to the functional requirements of the control centre, its support facilities, and its infrastructure. It can be viewed as a functional requirements specification for guiding the design and development of a control centre. It is important to note that it provides a total perspective of the operation of a control room, from human-machine interface issues and operations to administrative work. The Halden Project has developed control room philosophies for a number of different process units, including nuclear power plants and interim storage facilities for spent nuclear fuel.

An important motivation for the development of a control room philosophy is the early involvement of end-users in the design process. Since it is considered important that operational knowledge and experience within an organisation is fully utilised, operations-oriented documents are collected, which together with continuous dialogue with experts within operational and other departments, provide a basis from which a complete concept is developed. This process of information collection and refinement provides an opportunity for operational staff to provide feedback on ideas developed during the design process. VR-based design tools have been developed at the Halden Project that enable a three-dimensional model to be used to visualise and modify a design based on a control room philosophy. Software tools have also been developed to support the design documentation process.

A VR model can be used throughout the lifecycle of a control room, thus its useful life is not limited to the duration of the design process. After the initial conceptualisation and design stages of a project have been completed, the VR model can be used to plan construction, operations and maintenance, to define training sessions, as a basis for planning retrofits and, finally, to plan decommissioning. Thus the initial investment in a VR model for supporting the creation of a design prototype provides value later in the life of the control room.

2. THE DESIGN PROCESS OF A NUCLEAR CONTROL ROOM

Figure 1 describes the design process. The planning and concept development phase comprises of the conceptual design of the new control centre in terms of a philosophy. The design should be user requirement driven, not technology driven, which requires that great effort be devoted to the analysis phase. Continuous end-user participation and feedback is essential to foresee the end-users' design requirements. Close co-operation between the vendor and the client's operational department is therefore required throughout the design process. The design process is specified according to established standards and guidelines, i.e. primarily IEC 964, IEC 45A, NUREG-0711, and EPRI NP-3659.

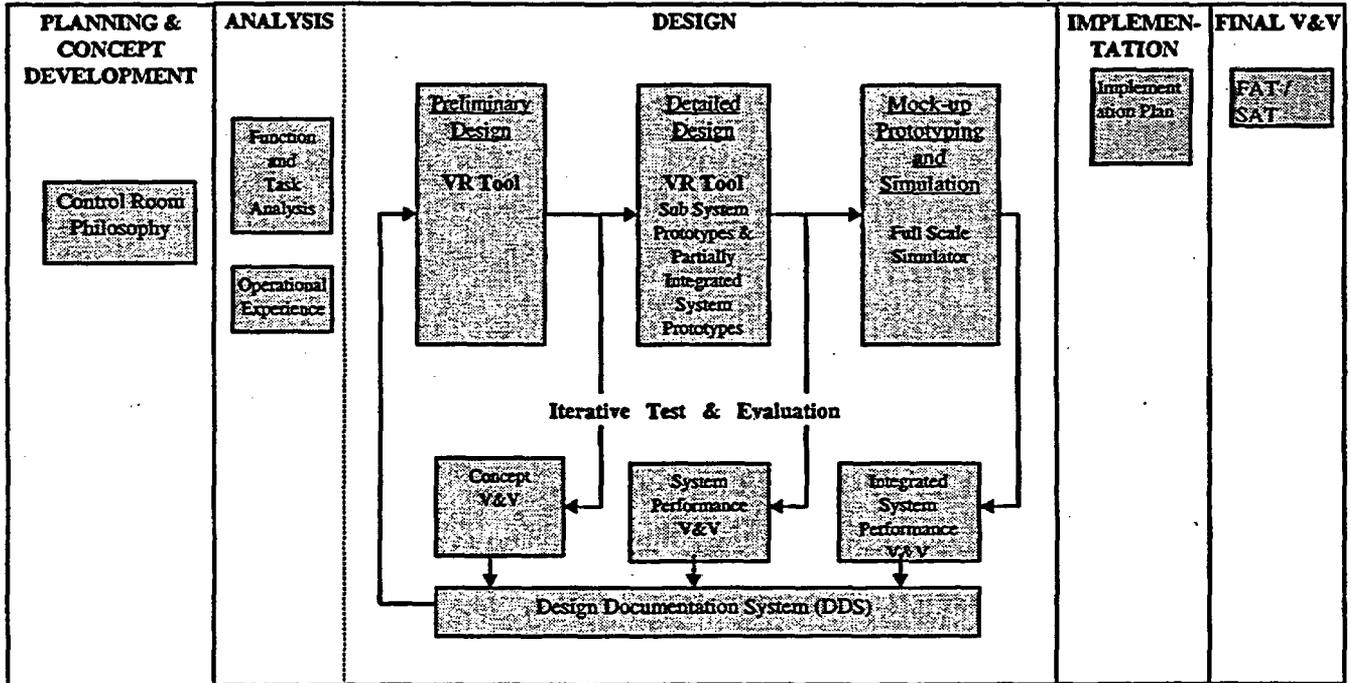


Figure 1: Different phases of the design process.

The outcome from the analysis work provides a detailed basis for the preliminary design, which consists of a human-machine interface (HMI) on workstations, a HMI on large screens, an alarm system design, and a control room layout. Prototyping starts as soon as possible in the design process and becomes a parallel activity; the first preliminary design specifications provide the basis for this work. In the ensuing detailed design phase, prototyping becomes increasingly advanced and development of a control room mock-up can eventually begin.

3. THE CONTROL ROOM PHILOSOPHY

The novelty with this approach to developing a control room philosophy is the concept of a total perspective for the entire control centre, in addition to the very early involvement of end users in the development work. By total perspective it is meant that all aspects of a control centre's functions are examined, including operational, administrative, and social aspects. The concept includes not only issues on human-machine interfaces, I&C or operational topics, but also administrative work and the infrastructure necessary to run a plant 24 hours a day, 365 days a year, including low power operation during outages. It is therefore necessary to consider the entire activity that is expected to take place within this environment. A brief summary of the content of a control room philosophy is provided in the following sections.

3.1 Roles of the Operators

The nature of the operators' roles in the main control room (MCR) depends on several factors. Three main role categories in the philosophy that can be listed for any of the operator positions in the MCR are *plant supervisor*, *plant operator* and *plant administrator*.

The supervisory role can be separated into two main categories: 1) passive monitoring, where the crew jointly supervise pre-defined plant parameters permanently presented as common information; and 2) active monitoring where the crew jointly, or individually, search for information to identify the nature of, and reasons for, an abnormal situation. The nature of this latter role is dependent upon the level of sophistication of the alarm system and available operator support systems. A philosophy typically promotes flexible alarm handling and presentation, where all detailed alarm information is available on request.

The plant operator role is primarily dependent upon the degree of automation, which in this context can be defined as the task distribution between the human operator and the control systems, with the goal to maintain the plant in a condition required by safety and operational goals.

The plant administrator role is primarily dominated by plant maintenance administration, supervision and record keeping; but also factors like logging and reporting, communication with other departments that relate to the overall plant operation and maintenance, and operator training.

3.2 STAFFING SIZE AND RESPONSIBILITIES

One condition that is often emphasised during the development of a control room philosophy is that the shift staffing must be maintained. A typical example of staffing size to be accommodated by the MCR is between six and nine operators:

- 1) One shift supervisor, with the overall responsibility for all activities related to the MCR functions.
- 2) Optionally, one assistant shift supervisor, who primarily provides support to the shift supervisor.

- 3) One reactor operator, trained for the control of the reactor and associated circuitry, as well as the balance of plant (BOP) side of the plant operation. The reactor operator can take over the BOP side when necessary, which means that s/he has full competence for the entire plant.
- 4) One turbine operator, trained for the control of the BOP side and all associated process and electrical circuits.
- 5-9) Station technicians, that work in the plant outside the MCR for most of the shift time. Their main tasks are to monitor and inspect component functions, and to perform all local manual operation on components. Workstations with process computer terminals are proposed in a new MCR philosophy for the station technicians.

The staffing size determines the baseline requirements for the number of workplaces, size, and layout of the MCR.

3.3 Control Centre Functions

This part of the control room philosophy identifies and describes the functional principles of the MCR. The requirements from the process assigned to the MCR provide the basis for the control centre functions necessary to operate the plant in all operating modes. On an overall level, these functions can be described as either operational or administrative. The operational functions comprise:

1. safety — what concerns the plant, the personnel, and the public;
2. operation — including both monitoring and process control actions;
3. co-ordination of all operational staff.

Administrative functions are all functions that cannot be related to operational activities, but are performed by the control centre staff. Roughly they can be divided into plant maintenance-related work (e.g. during outages); reporting and data logging; staff training; communication with operational department and other departments; and production, updating, and reviewing of administrative documentation e.g. procedures, technical system descriptions etc.

3.4 A Functional Control Centre Layout

A functional control centre layout is developed from the top by defining the infrastructure and the environment around the MCR. The most important support functions are: conference room, control room office, work permit management office, entrance to the control room, and an emergency control room (ECR) located at a physically different location than the MCR. The MCR also needs storage facilities, a kitchen, and so forth, but these areas are not specifically addressed in this paper.

3.4.1 Emergency Control Room (ECR)

The ECR is designed for use when the MCR is no longer accessible due to fire, sabotage, contamination, etc. The ECR must therefore include all necessary I&C equipment to be able to shut the plant down, firstly a hot shutdown and secondly a cold shutdown. The layout of the ECR should

principally follow the MCR to simplify the transition of the plant operation when leaving the MCR and using the ECR. This will minimise the operator efforts when the transition has to take place, which can be a very stressful event. The ECR will never replace the MCR, but should be equipped to manage certain functions in very specific situations. The ECR should include the same type of computerised operator interface as the MCR for situations where the computer systems are still accessible, which will support normal working procedures. In addition, the ECR should include an analogue stand-by system completely outside the digitised system in case of a total computer system blackout.

3.4.2 Conference Room And Office Area

An important feature for administrative work is the close location of a conference room for daily meetings where the operating crew participates. Typical examples are meetings between the shift supervisor and the operational management, maintenance departments, etc. An office is usually also required.

Both a conference room and an office could, for example, be located at the back of a MCR, with glass walls between the rooms and the MCR to maintain a good overview of what is happening in the MCR. The rooms should be equipped with computerised operator workstations on the process computer systems, but with the restriction that no control action can be performed. A special entrance should exist to the conference room so that no visitors disturb work in the MCR.

3.4.3 Work Permit Management Office

This area should be used for all planning and administration of work permits for maintenance work at the plant. It should contain office space for 8-9 people (depending on the type of control room), equipped with necessary equipment. This includes personal computers, operator workstations without control access, in addition to all important technical documentation. Different reception desks can be included, directed towards both the MCR as well as a radiation protection office, and restricted / non-restricted areas for radiation. This is for verbal and written communication between this office and the MCR / maintenance staff.

3.4.4 Entrance To MCR

There is a safety aspect related to the location of the MCR entrance. It should be possible to place the entrance within the sight line from the MCR workstations to give the operators a good, immediate, overview of who is entering the control centre area.

3.5 Workstations and their Ergonomic Layouts

The main equipment at the operator workstations typically includes process display monitors, a central TV surveillance system, telecommunication devices, and a personal computer for administrative work. The philosophy supports the idea of spreading process information to different receivers who have an interest in the process state, e.g. the operational management, outage planning, technicians during local operations or inspections, etc. This can be done by installing operator stations on a computer network at different locations at the plant. Of course, no control actions be performed from these units.

Each workstation in the MCR should, in addition to monitoring and control capabilities, have sufficient space for administrative work (desk area), good communication opportunities within the MCR as well as with staff outside it, and space for all necessary documentation. The workstation should be ergonomically designed with, for example, the possibility to adjust the height and local lighting.

Extra space is generally desirable in the MCR and at the workstations for staff that are undergoing on-the-job-training in parallel. This is common for all units and for all job positions.

3.6 I&C Functions

These are examples of main parameters for an I&C functions' philosophy. To comply with the situation when the crew works as one team, the philosophy states that common information important to all crew members has to be visible from all locations within the MCR, and shall be provided on centrally located large screens. The screens should present partly pre-defined information (e.g. plant overviews, alarm information, etc.) and partly ad hoc information as selected from any of the workstations. With this, the philosophy will provide flexibility concerning the large screens with both a pre-defined overview permanently displayed as well as selectable displays dependent on the situation. These can be presented simultaneously. Furthermore, since the shift team works either as a team, a group within a team, or as an individual, the philosophy states that all available information shall be accessible on all MCR workstations. Process control is possible from all operator workstations. This provides individual workstations equipped for the supervisory function; the reactor monitoring and control; and the balance of plant operations.

3.7 One Consistent Information Presentation System for all I&C and Support Systems

Information presentation should, in the best possible manner, support the operators in their primary tasks and goals for a safe and economically sound operation of the plant. It is important to have a flexible system where parts can be replaced without affecting the user interface. The user interface has to be internally consistent and unified, independent of the kind of computer system that is used. New operator support systems that will be implemented in the future should be able to be integrated with the process computer system.

It is important that the computer system does not require the user to know how this system is technically functioning to be efficiently used. This can be achieved by an efficient design of the user interface. It is also important that the computer system does not have limitations in terms of what kind of graphics etc. that can be displayed. Other e.g. some graphical forms of representations of the process can be more efficient to use than present formats. Examples of principles from a philosophy are: the display formats shall be designed in such a way that the need for information gathering from several formats is minimised; display formats shall be logically presented to the operator; the user shall himself / herself be able to control amount and complexity of the information presented.

3.8 Human-Computer Interaction and Process Control Needs

All process control input in a modern computerised control room is through the computer, which in turn controls the actual component. One of the most important challenges is to develop a computer interface to the process that requires a very low level of skills for interaction. The operators must feel that they are interacting with the process not the computer.

This section of the control room philosophy defines the principles and the most important corresponding guidelines for computer interaction and process control. It is structured in two parts: 1) interaction between operators and the computer interface (i.e. display suite navigation), and 2) the interaction between operators and the process (i.e. process monitoring and control).

3.9 Alarm System Requirements

Computerised alarm systems provide the capability to prioritise, suppress, and filter alarm signals to suit the event. The number of alarms are in this way reduced by filtering consequence alarms that do not provide any additional information to the situation but instead increase the operators workload and disturb diagnosis of the situation by requiring the acknowledgement of these alarms. The alarm philosophy identifies the methods that can be used to prioritise alarms to improve diagnosis, and to find and act on the root cause of the event. Principles are defined to structure and integrate the alarms in the process displays and the alarm displays on the terminals and the large screens. A special section of an alarm philosophy deals with presentation and integration of high level alarms from separate advanced alarm systems, such as alarms from an early fault detection system and diagnosis systems.

Examples of principles from an alarm philosophy are: The "black screen" approach shall be used, where alarms are only presented to attract the operator's attention to indicate faults that require an operator to intervene. Alarm presentation shall adhere to the principles for the information presentation system. Alarms shall be structured in a causal tree format to identify the connections between causal alarms and consequence alarms, etc.

4. VIRTUAL PROTOTYPING OF A CONTROL ROOM PHILOSOPHY

The VR-based design tool developed at the Halden Project is a central component of this structured approach to developing a control room design specification. The main motivations for the VR design tool are the enabling of early participation of end-users and the removal of the need for a physical mock-up. In particular, it enables the actual control room operators, who will eventually use the new control room, to participate in the design process. It is believed that experienced operators, in continuous dialogue with human factors engineers and system vendors, have the greatest potential to develop an optimal control room design at a sufficient levels of detail.

Control room operators are pragmatic experts on what information should be available in a control room and on work procedures in normal, disturbed and outage operations. However, because operators are not trained design engineers with broad knowledge of alternative solutions, they benefit from advice and guidance throughout the design process. Since operators are usually not experts at

expressing design solutions using traditional computer-aided design (CAD) tools, the VR tool described in this paper was developed. The VR tool enables operators to design new control rooms by selecting objects from libraries and positioning them in a room. Objects such as furniture, safety panels, computer monitors, and walls, can be moved, scaled and reshaped to visualise a design idea. The operators do not model the actual objects themselves, but can request that a modelling expert add new objects to the object library.

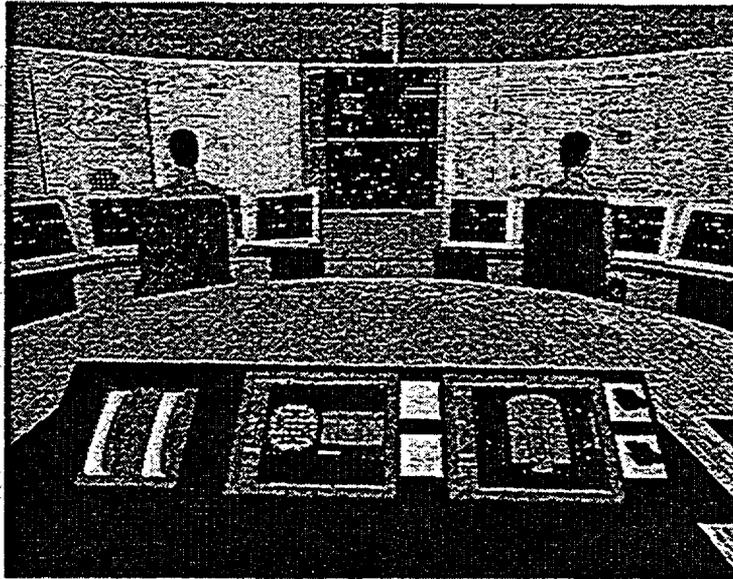


Figure 2: The shift supervisor's view from his/her workplace

A significant part of the design process is a human factors test and evaluation process. This is an iterative process from the conceptual design stage of the project to the system testing stage. Tools have been developed at the Halden Project to assist in the verification and validation (V&V) of control room designs, using a VR model to verify designs against design specifications. International standards and guidelines for the nuclear industry are used as a basis for the V&V tests supported by these tools. In the future, it is intended that the tools will be enhanced to support the validation of functional requirements (using walkthroughs), operational procedures and training programs.

Typical V&V design tests include evaluating reach and posture, checking sight lines and examining ergonomics. For example, figure 2 shows a shift supervisor's view from his/her workplace. This view can be evaluated using the tests provided by the V&V tool. In this example, the reactor and turbine operators' monitors obscure some instruments on the panels when the shift supervisor is in a sitting position. The monitors cannot be lowered or angled further, according to human factors recommendation. There are two available options. Either the shift supervisor's workspace can be raised a step higher in the control room or a step can be placed under the panels. Figure 3a shows a mannequin being used to evaluate the effect of adding a step to the panels. Figure 3b shows the mannequin being used to evaluate the effect the step has on reach and posture.

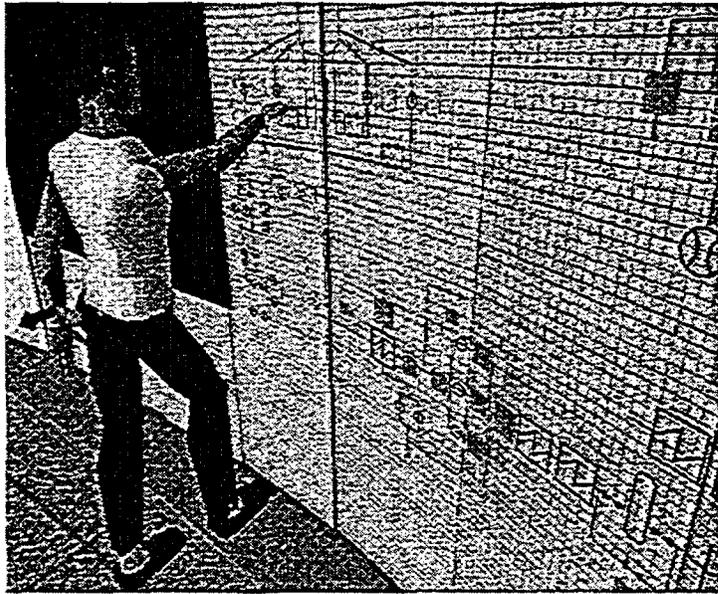


Figure 3a: Mannequin used to evaluate step in front of a panel

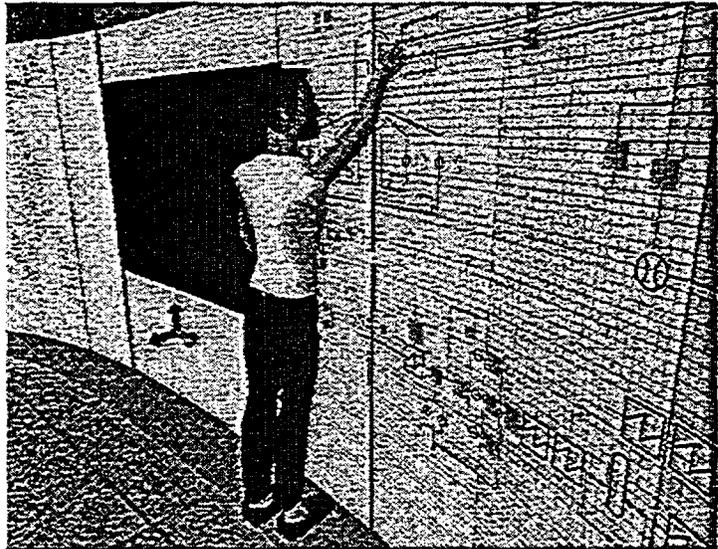


Figure 3b: Mannequin used to evaluate reach and posture

[Note that a major distinction exists between design process verification and validation (as described above) and the final design implementation verification (factory and site acceptance tests, FAT/SAT)].

5. USING THE VR TOOLS

In practice, a VR laboratory is established at the site where the control room operators are located. The lab typically consists of a graphics workstation with appropriate software installed and a projection system, to facilitate group discussion. 3D input devices and LCD shutter glasses for stereoscopic viewing can also be used.

The members of the design team are trained to use the VR tools and the design documentation system (DDS). The VR lab is used as a regular meeting place for all participants in the design process. The design team follows the guidelines in a control room philosophy to develop a control room prototype using a structured approach. A typical working day is a combination of design and documentation sessions.

The initial starting point is a model of a room (walls, ceiling, and floor). The designers can then add windows and doors and add or remove walls. The primary objects, such as safety panels, large screens and workplaces can then be placed in the model. In the case of a retrofit, the initial model typically depicts the layout of the existing control room, otherwise it shows an initial idea for a layout that the team of designers agree upon as a starting point. At this stage the virtual control room provides a context within which the designers can focus on specific areas of a control room philosophy, continuously refining the initial model until a final design is reached.

The initial focus is typically on the layout of instruments, symbols, and mimic on instrumentation panels, then the design of operator workplaces, the layout of the control room, and, finally, the layout of the entire control room suite and facilities. The design of process displays can be done in parallel with the other design activities.

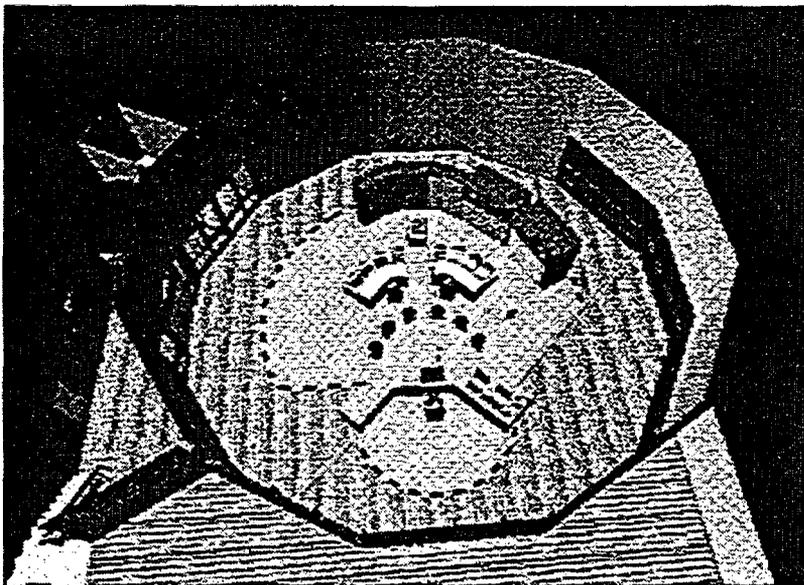


Figure 4: An initial model of a control room with the main objects

Figures 4 to 7 show some of the VR tools in use. Figure 4 shows a model of an existing building, within which a number of objects have been positioned by the design team. The designers add new objects to the model by picking objects from an object library and dragging them into position. Figures 5 and 8 show the object library in action.

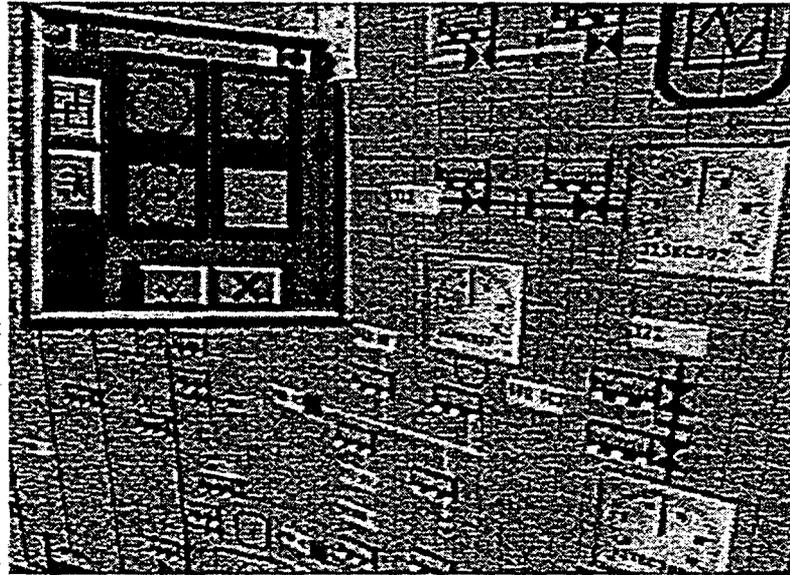


Figure 5: Object library showing instruments and symbols for a safety panel

In addition to supporting the creation of furniture and other relatively large objects, the object library is also used when designing the layout of analogue safety panels. The designer selects a panel, then picks, from the library, instruments, lines, symbols, etc., which are automatically attached to the surface of the panel. The designer can then drag these objects around the surface of the panel into position. When a mosaic system is used, objects 'snap' into position. The objects themselves are models of the components available from the selected vendor. Objects can be grouped together, and entire groups of mosaic components can be copied or moved around the surface of a panel, or even be copied to other panels. The designer has great flexibility to experiment with the positioning of components in order to find an optimal layout. Text labels can be added to label objects and alarm tiles as appropriate, and a standardised colour palette can be used to modify the colours of lines and symbols. A final panel layout can be exported from the VR model in a CAD format to print technical drawings of the panel designs. By connecting the virtual panels to a simulator, it is possible to evaluate the behaviour of the panels.

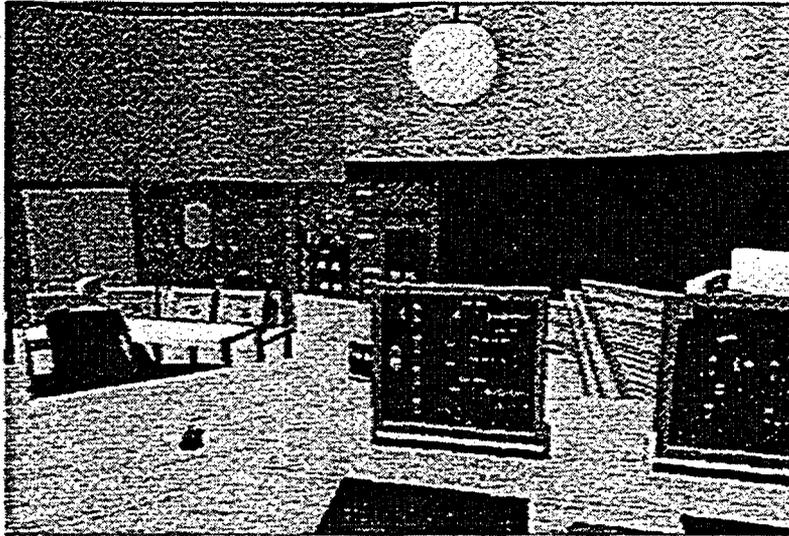


Figure 6: Evaluation of an operator workplace

The object library contains a number of different designs of desks and chairs that can be used to design work places. Additional furniture can be added to the object library as necessary. The designers can interactively create new work places in the model and evaluate them. The VR tool enables the designers to evaluate sight lines and measure distances to ensure that information on computer monitors, large displays, and safety panels is readable and that visual communication between operators is sufficient. Since the VR tool enables designers to position themselves anywhere within the virtual control room, it is easy to check that furniture and people do not obstruct vital information.



Figure 7: Control room layout evaluation



Figure 8: Object library with furniture and other equipment to be placed in the control room

Similarly the design of the supporting infrastructure is developed. Offices with access to the MCR are positioned, windows added where appropriate, and facilities such as a conference room, rest area, kitchen and visitors' gallery can be added, in accordance with the control room philosophy. For example a control room philosophy might recommend that a conference room should have entrances to the MCR and to a secure area outside the MCR, so that it is possible to enter the conference room without passing through the MCR.

6. THE DESIGN DOCUMENTATION SYSTEM

As the design develops, it needs to be documented, and the documentation should include all tests and evaluations. Some of the design documents form the basis of the design specifications for a vendor, while other are used to demonstrate that a high quality approach to the design development has been followed, with many iterative tests and evaluations (as required by licensing authorities). To assist the designers/operators in this task, a computer-based Design Documentation System (DDS) has been developed and integrated with the VR tool. Pages in the DDS are used to store descriptions of design solutions, with the ability to immediately retrieve and view relevant VR models, providing an advanced version control mechanism. In addition to support for version control, the DDS also comprises of the V&V tool described briefly in section 4.

Design documentation is written as the design changes. Information in the DDS guides the designers to document design arguments, references to standards and exceptions from philosophy guidelines or standards for all objects in the control room. When the design is frozen, reports for different purposes are generated by the DDS. These reports can serve as the actual design specification, including pictures from the virtual control room that illustrate the design. All reports from the DDS can be edited using a word-processor package. A particularly useful type of report that the DDS can produce is a "design change status report", which is regularly printed to keep track of suggested changes to the

developing design. Immediately before the design is frozen, this report serves as a check-list of design ideas which may not have been sufficiently followed up by the designers.

An important feature of the DDS is that it supports the easy handling and browsing of documents. Documents can be copied, moved, and deleted, and new documents can be created in a document hierarchy with headings specified by the designer. The DDS supports a top-down approach to the design documentation, so an overall design philosophy is documented at the highest level, followed by guidelines that support the philosophy, followed by defined standards. The level below the system specific standards is the actual design. Figure 9 shows an example of a DDS document structure. During the design process, it is likely that compromises will have to be made due to existing constructions and limitations, which force the designer to deviate from defined standards. These exceptions are documented in the DDS together with other design arguments that originate from operational experiences or good practice.

Key features of the DDS include:

1. **Structured top-down approach to the design.** All documents regarding design philosophy, concept, and general guidelines and standards for each part of the design are available at a high level in the DDS. Only deviations from these are documented lower down in the hierarchy, to avoid unnecessary amounts of documentation and to be able to trace all deviations.
2. **The design documentation, such as design arguments, standards, guidelines, and references, is written into the DDS at the time when a design or redesign decision is being made.** This approach avoids time-consuming periods dedicated to design documentation, as documentation is created online as each problem/issue is discussed and the design arguments are identified.
3. **Design documentation can be written directly into the DDS using forms or can be imported from Microsoft Word (or some other organisation-standard document format).** Documents in the DDS can be exported for reading or manipulating using other software packages.
4. **All design documentation can be accessed and viewed over an Intranet (or the Internet, if required) by anyone that has been supplied with a login-in account and password by the DDS administrator.** Thus, end-users of the design and other people that are interested or have input to the design progress, have access to the design descriptions that are of interest to them and give feedback to the design engineers.

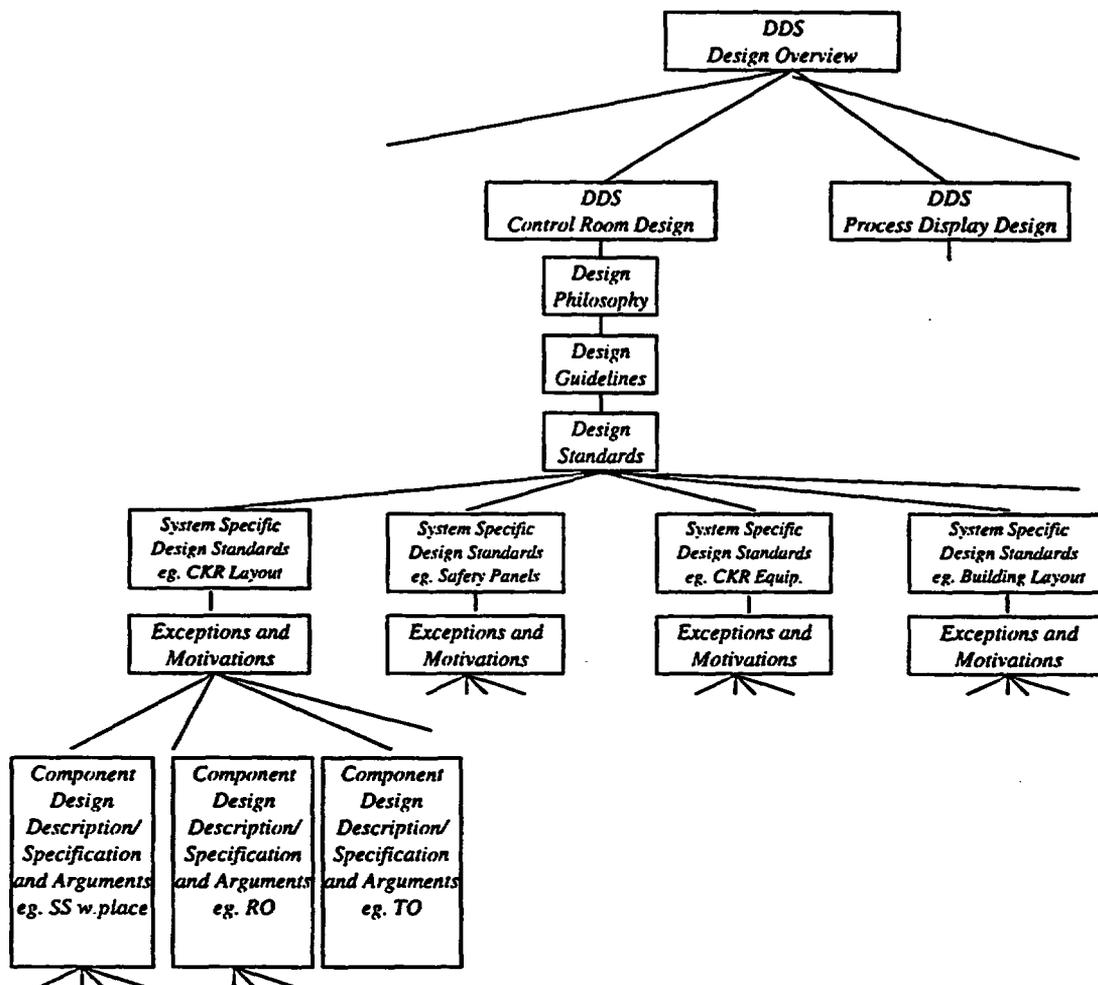


Figure 9: Example of top-down design documentation structure in DDS.

7. SUMMARY

The application of VR in combination with an advanced control room philosophy provides a complete approach for the design of control rooms. The techniques and tools described in this paper have been used in a number of projects in both the oil and nuclear industries. Design development time is believed to be shortened and the design acceptance of operators appears to be high. Virtual control rooms have been found to be good replacements for physical mock-ups. The VR tools described are under continuous development, with new features being added to assist design teams.

Full Range Signal Validation of PWR Plant Data And Fast Transient Classification Applied In Alarm Handling Using Neuro-Fuzzy Models

P.F. Fantoni, D. Roverso, F. Øwre
OECD Halden Reactor Project, P.O. Box 173, N-1751 Halden, Norway
Tel: +4769212200 – Fax: +4769212201
E-mail: Fridtjov.Owre@hrp.no

Abstract

The surveillance and control of any industrial plant is based on the readings of a set of sensors. Their reliable functioning is essential since the output from the sensors provide the only objective information about the state of the process. The signal validation task is to confirm whether the sensors are functioning properly.

Real-time process signal validation is an application field where the use of fuzzy logic and artificial neural networks can improve the diagnosis of faulty sensors or drift in sensor readings in a robust and reliable way.

The present work describes the transient and steady state on-line validation method of plant process signals using artificial neural nets (ANN) and fuzzy logic pattern recognition. This method has been developed at the OECD Halden Reactor Project and tested on simulated scenarios covering the whole range of PWR operational conditions provided by Electricite' De France (EDF) and the Centre D'Etudes De Cadarache (CEA) in France.

Events and faults in nuclear power plants can set off transients, which subsequently can activate a large number of alarms presented in a rapid sequence to the operators. A robust method to suppress less important alarms has been sought for a long time. Starting from the work outlined above the Halden Project has developed a ANN based system performing a fast classification of the occurring transient, then providing this information as input to an alarm handling system which again applies this information to perform event-driven alarm suppression. The paper reports on the first phase of this project including a description of the method and the prototype system.

1 Introduction

The operation of each industrial plant is based on the readings of a set of sensors. Their reliable functioning is essential as the output of sensors provides the only objective information of the process. The task of the signal validation is to confirm whether the sensors are functioning properly. Signal validation must be enough robust to multiple sensor faults as well. This requirement is crucial especially in case of an accident when the abnormal changes of the process together with possible severe damage of the sensors can occur.

The present work describes the transient and steady state on-line validation method of the plant process signals using artificial neural networks (ANN) and fuzzy logic pattern recognition. The use of ANNs for signal validation has several advantages. The most important are - it is not necessary to define the physical model of the monitored process and properly trained ANNs are less sensitive to the measurement noise than the model-based techniques.

This paper represents the continuation of the work at the OECD Halden Reactor Project¹. The signal validation model is based on the set of the ANNs, each driven by a pattern recognition algorithm. This classifier based on the fuzzy and possibilistic clustering technique identifies the incoming signal pattern (a snapshot of process signals) as a member of one particular cluster from a set of clusters. They are recognized to cover the entire operating range represented by the possible combinations of steady state and transient values. Each cluster is associated with one ANN previously trained only with data belonging to this cluster. During the operation the classifier provides an automatic switching mechanism to allow the best-tuned ANN to be used. The maximum membership grade of the sample in the particular cluster and the maximum signal mismatch in the neural network module input into the Mamdani type fuzzy model to estimate the reliability level of the validation. This model has been developed and tested on simulated scenarios covering the whole range of PWR operational conditions² provided by Electricite' de France and the Centre D'Etudes De Cadarache. It does not require any special and additional type of measurement or equipment (common e.g. in noise diagnostics) as it utilizes only the standard measurements available in the plant.

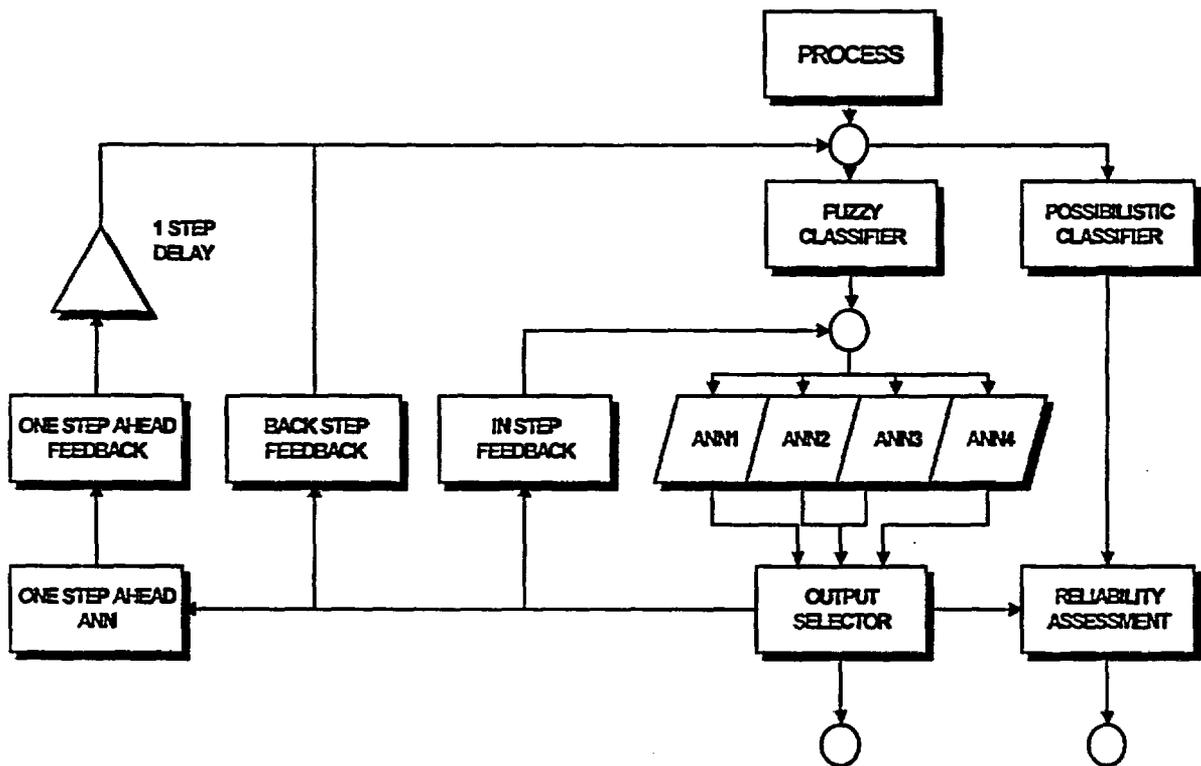


Figure 1 Neuro-Fuzzy model block diagram

2 Fuzzy classification

Let $\bar{x} = [x_1, x_2, \dots, x_N]^T$ a vector in \mathcal{R}^N representing an input dataset. The N components are correlated process signals that constitute a snapshot of the monitored process at a given time. Given

$X = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_p)$ the $N \times P$ matrix of P patterns covering the \mathfrak{R}^N operating region, the basic idea is to split this region in Q fuzzy clusters and derive a mapping function which assign each pattern $\bar{x}_i | i = 1 \dots P$ to each cluster $C_k | k = 1 \dots Q$ at *some degree*. This transformation is expressed by the following equation:

$$\bar{x}_i \Rightarrow \{u_{1k}, u_{2k}, \dots, u_{Qk}\} \quad (1)$$

And

$$u_{ik} \in [0, 1], i = 1 \dots Q, k = 1 \dots P \quad (2)$$

Where u_{ik} is the membership grade of the pattern \bar{x}_i in cluster C_k . In pattern recognition the Q clusters are identified by prototype patterns, which in the case of spherical or ellipsoidal clusters are also called centroids, so that the representation of a fuzzy classifier for a given X ($N \times P$) matrix dataset with Q clusters is completely defined by:

$$B = (\bar{\beta}_1, \bar{\beta}_2, \dots, \bar{\beta}_Q) \quad (3)$$

$$U = (\bar{u}_1, \bar{u}_2, \dots, \bar{u}_Q)^T \quad (4)$$

$$\bar{\beta}_m = [x_1^{C_m}, x_2^{C_m}, \dots, x_N^{C_m}]^T \quad (5)$$

$$\bar{u}_m = [u_{m1}, u_{m2}, \dots, u_{mP}]^T \quad (6)$$

Where B is the $N \times Q$ matrix of the cluster prototypes and U is the $Q \times P$ matrix of the membership grades of X , also called the *fuzzy C-partition*.

The fuzzy partition problem, as expressed in (3), (4), (5) and (6) can be solved with the minimization of an objective function¹⁶, which can be written as:

$$J(B, U, X) = \sum_{i=1}^Q \sum_{j=1}^P (u_{ij})^m \Delta^2(\bar{x}_j, \bar{\beta}_i) \quad (7)$$

Where $m \in [1, \infty]$ is called the fuzzifier parameter and Δ is a function representing the distance between two vectors. When $m = 1$ the classifier is crisp and when $m \gg 1$ fuzziness is maximized. $m = 2$ is the recommended value, for most applications.

The choice of the Δ function depends by the expected shape of the clusters. If the Euclidean distance is used, which is the right choice for spherical clusters, the resulting algorithm is the popular *Fuzzy C-means* algorithm. In this application clusters with different shapes and sizes are expected, so that Euclidean distances would not work well. To take care of the not uniform distribution of the patterns in the dataset, the GK algorithm, from Gustafson and Keller³, has been used. Here the distance function is expressed as:

$$\Delta_{i,j}^2 = (\det C_i)^{\frac{1}{N}} (x_j - \beta_i)^T C_i^{-1} (x_j - \beta_i) \quad (8)$$

Where C_i is the *fuzzy-covariance* matrix for cluster I, defined as:

$$\frac{1}{\sum_{j=1}^P (u_{ij})^m} \sum_{j=1}^P (u_{ij})^m (x_j - \beta_i)(x_j - \beta_i)^T \quad (9)$$

In fuzzy clustering U must satisfy the following three conditions:

$$\sum_{i=1}^Q u_{ik} = 1, k = 1 \dots P \quad (10)$$

$$u_{ik} \in [0,1], i = 1 \dots Q, k = 1 \dots P \quad (11)$$

$$0 < \sum_{j=1}^P u_{ij} < P, i = 1 \dots Q \quad (12)$$

Condition (10) reflects the probabilistic requirement that the total probability for an input dataset pattern to belong to *any* cluster is 1. In other words, patterns not reflecting any of the identified cluster prototypes are classified and assigned to the *relatively* most probable cluster, only because of the implicit certainty that *all* the patterns belong to the established partition. There can be uncertainty (or fuzziness) on *where* the incoming pattern could be assigned, but no uncertainty on *if* it can be assigned somewhere. When this methodology is applied to signal validation applications, a number of problems may arise:

- Lack of robustness against noisy data. There is no compensation for the noise in the calculation of B and U .
- It is not able to say "I do not know", also when this would be the best answer. An incoming pattern might be given a high grade of membership in a cluster, even if it is far away from all the centroids, only because it is *relatively* closer to one specific cluster.

Relaxation of requirement (10) leads to a possibilistic approach, that results in a possible solution of the two above mentioned limitations.

A possibilistic classifier initially learns a dataset X of pattern samples (in other words it calculates B and U). During this process, the model increases its robustness to noisy data and many patterns in X could be discarded as not representative of any developing cluster. When new patterns are examined, the possibilistic model evaluates in which cluster or clusters the incoming pattern could be *possibly* assigned, if any.

Following Krishnapuram and Keller's work¹⁷, minimization of the objective function (7), without the constraint in (10), results in the following equations for U and B :

$$u_{ij} = \frac{1}{1 + \left(\frac{\Delta^2(\bar{x}_j, \hat{\beta}_i)}{\eta_i} \right)^{\frac{1}{m-1}}} \quad (13)$$

Where $m > 1$ and

$$\beta_i = \frac{1}{\sum_{j=1}^P (u_{ij})^m} \sum_{j=1}^P (u_{ij})^m \bar{x}_j \quad (14)$$

Where η_i is computed by:

$$\eta_i = (\det C_i)^{\frac{1}{N}} \quad (15)$$

The step-by-step procedure used to develop the fuzzy and possibilistic classifiers can be summarized as follows:

- Given a set of samples X , compute an initial set of cluster centroids using the ISODATA algorithm, that has been chosen because it automatically optimizes the number of required clusters.
- Initialize the elements of the partition matrix U with crisp values (0 or 1), using ISODATA. Then run the GK algorithm, which produces the fuzzy classifier.
- Use the updated matrix U and B , from the previous step, to start the iterative process as shown in eqs. (13) and (14) to arrive to a possibilistic partition.

3 The artificial neural networks module

Sample data in dataset X are collected in Q training datasets, to be used for training Q supervised neural networks. Each pattern in X is assigned to one or more training set according to the fuzzy partition, as long as its possibilistic index in U is above a threshold value h in one or more identified clusters, with $h = (0.5, 1)$. The role of the threshold parameter h is twofold:

- sample patterns not adequately represented in any cluster are discarded, so that they have no influence on the network weights calculation.
- sample patterns *possibly* represented in many clusters (responsible of the above mentioned boundary problem) are used in the training set of many corresponding networks.

The network architecture used in this work is a five layers (three hidden layers), feedforward structure trained with the backpropagation algorithm. For better performance, a momentum term and an adjustable learning coefficient have been used. The hidden layers use hyperbolic tangent transfer functions, while the output layer is linear. This architecture has been proved¹ to be more robust to process noise and sensor faults.

The input to the ANN's is not limited to the current pattern. To capture the process dynamics, a number of past values of the time series are used, together with the current ones, so that the total number of input nodes in each ANN is $N \times R$, where N is the number of signals and R the number of past values used. In this work, samples at $t-0$, $t-2$, $t-5$, $t-13$, $t-34$ and $t-89$ have been used for each signal. This allows to consider both short and long time constant effects in the process dynamics.

The three feedback loops in Fig. 1 are used during the recall (on-line validation).

The *in-step feedback*, in case of mismatch in one or more signals, re-evaluates the corrected pattern to get better estimates in the not affected channels. This feedback is triggered only if the signal mismatch is estimated to be above two standard deviations, to avoid instabilities.

The *back-step feedback* corrects mismatching signals for the future evaluations, when those values will be used as past values, as mentioned before. This also triggered by the same threshold value.

Finally, the *one-step-ahead feedback* monitors the next coming pattern for possible large deviations. Large deviations have a negative effect on the overall performance, because they lead to false classifications with the results that not optimal ANN's are triggered for recall, only as a consequence of a large deviation in one or more channels. Large deviations from the expected values are corrected before the classification, resulting in a much more accurate and stable validation. In Fig. 1, a predictive ANN calculates one-step-ahead expected values, but tests have shown that the current values (at time t) can be used as a rough estimation of the values at time t+1.

The recall strategy does not make use of the above threshold parameter to trigger one or more of the specialized networks. The algorithm used here, applicable only to signal validation processes, applies the concept of *presumption of no sensor fault, if possible*. This concept is based upon the well-tested hypothesis¹ that when a neural network confirms the sensor input (no fault condition) and the relative cluster membership grade is high, the network output is reliable.

This leads to the following recall strategy:

- For each process sample: get the most representative cluster, which is the one with the highest membership grade u_l in U
- Recall the output using the neural network associated to this cluster
- Calculate the maximum absolute deviation, as follows:

$$err1 = \max |s_j - o_j| \times \text{sgn}(s_j - o_j) \quad j = 1, \dots, N \quad (16)$$

Where s_j and o_j are the j -th network input and output values for a process pattern

- If $err1$ is very low, accept the result, otherwise recall the pattern using also the network with the second highest grade of membership, $u2$
- Calculate the following weighted error:

$$werr = \frac{err1 \times u1 + err2 \times u2}{u1 + u2} \quad (17)$$

Where $err2$ is the maximum error with the second network. Now if:

$$abs(werr) > abs(err1) \quad (18)$$

Accept the output from the first network, otherwise calculate a weighted output from:

$$wout = \frac{out1 \times u1 + out2 \times u2}{u1 + u2} \quad (19)$$

Where $out1$ and $out2$ are the vector outputs from the two networks.

4 The reliability assessment module

In a previous work⁴ we tried to solve the reliability problem connected to the use of neural networks, using a Radial Basis Function network associated to a crisp pattern classifier. This work extends the idea, exploiting the unique features of a possibilistic classifier.

The possibilistic cluster membership has an important role in the final decision whether the network output can be considered reliable or not. A high membership grade in one or two clusters increases our confidence that the data sample is contained in the training volume of one or two neural networks, so that they will be able to recall the output with low estimation error. On the other side, a low membership value in *all* the clusters is a clear warning that no network has been trained to recall such a pattern. Note that using fuzzy clustering techniques, it would not be possible to have neither *low* values in all the clusters, nor *high* values in more than one.

In this work, the reliability function is realized through a fuzzy model (fig. 2), where the input is the maximum membership grade of the sample and the maximum signal mismatch in the neural network module, while the output is the reliability membership grade in three fuzzy sets assessing at what extent the reliability factor can be considered *high*, *medium* or *low*. This fuzzy model applies *Mamdani-type* implication rules, with the following fuzzy rules:

IF *max-grade* is *low* AND *max-mismatch* is not *low*
 THEN *rel-grade* is *low*
 IF *max-grade* is *high*
 THEN *rel-grade* is *high*
 IF *max-grade* is *medium*
 AND *max-mismatch* is *medium*
 THEN *rel-grade* is *medium*

For each data sample presented to the system, the three fuzzy rules are *fired* at different degree, resulting in three different membership values for *rel-grade* in the three fuzzy sets *high*, *medium* and *low*. These values can give a clear idea about the accuracy of the network output.

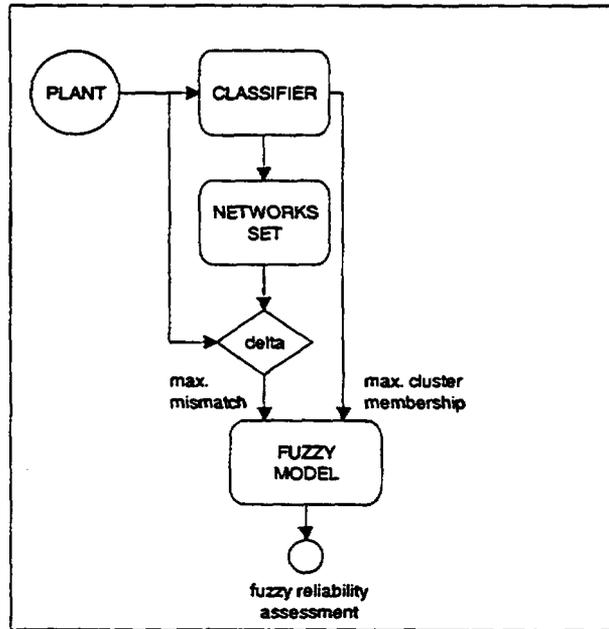


Figure 2. The reliability fuzzy model

5 Results of performance tests

Training and testing data from a French PWR simulator were used to perform the validation of 14 process signals (e.g. reactor power, coolant temperature, control rod position, boron concentration, feedwater flow, steam flow etc.). The training data were from normal *operation* conditions, varying from 25% to over 100% of reactor power. To improve the behavior of ANNs during the recall phase, the technique of the so called robust training was used by adding the noise of 2% and random faults of 20% of the range for each input signal one at a time, while retaining the noise free values for the desired output. During training, the generalization ability of ANNs was tested on the cross-validation data set – 10% of patterns put aside from the training set. Fuzzy classification resulted in 15 clusters and 15 corresponding ANNs.

The EDF/CEA tests contained five different cases of variable plant conditions and simulated one or more sensor faults. The nature and location of these faults as prepared by CEA at Cadarache were not known in advance, so that no model fitting to these blind test data was possible. Different kinds of multiple faults were simulated - noise superposed on the signal, bias and drift in the signal, as can be seen from Fig.3-5.

In the pictures the actual values of signals are the thin lines while the predicted values are thick solid lines. The error bands in the mismatch plots are calculated by the model according to the expected error

of prediction for each particular cluster, calculated in advance during the training. The error bands should be interpreted as follows:

- first band (dashed) : It is set at two standard deviations of the expected error. Exceeding this band is considered as the first warning, especially if the situation persists.
- second band (solid) : It is set to four standard deviations. Exceeding this band is considered a defined alarm condition
-

During the *normal operation scenarios*, which the model was trained for, all faults were recognized *correctly*, with the reliability level of *high* or *medium*. The alarm was triggered either instantly (bias, noise) or after few samples (drifts) as soon as the mismatch between the signal and the estimated values exceeded the second error band.

Test 3

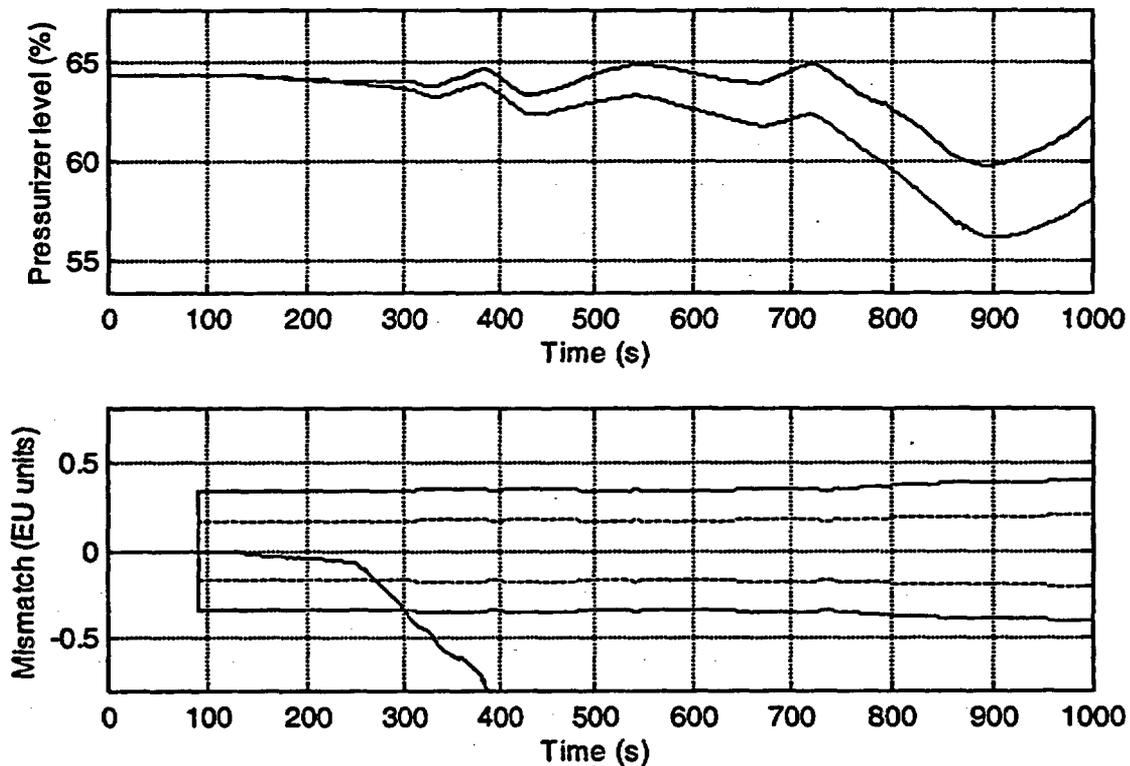


Figure 3 Pressurizer level drift

Test 2

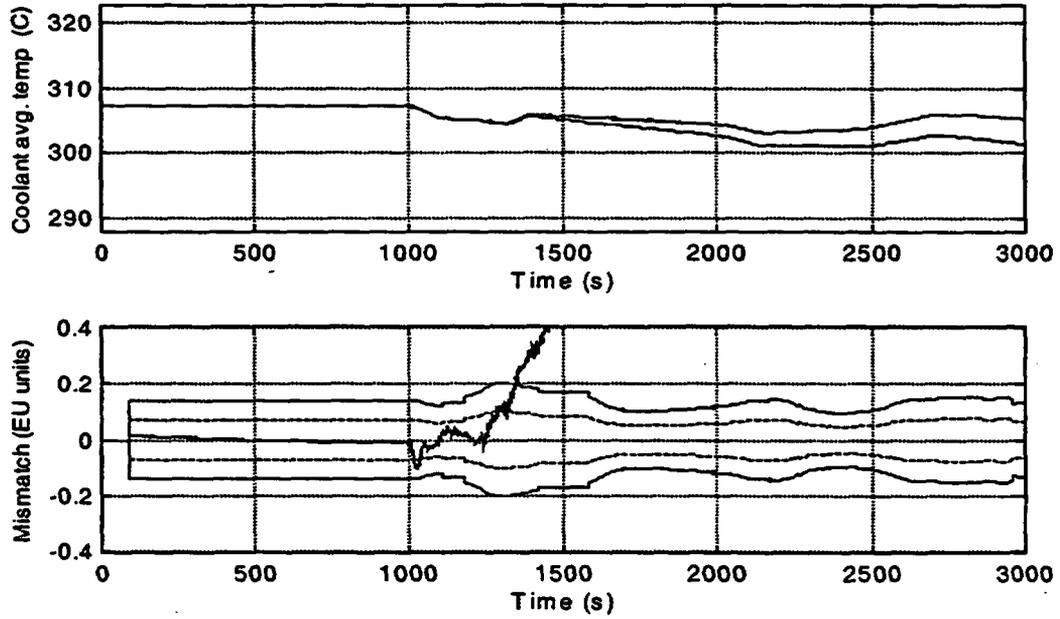


Figure 4. Core coolant temperature drift in multi-failure scenario

Test 2

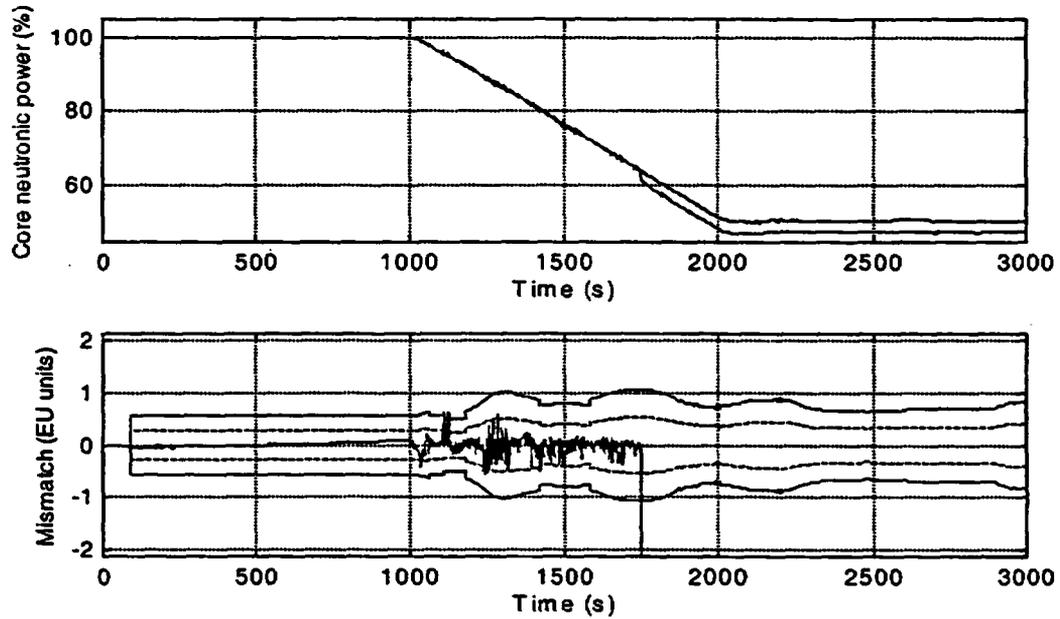


Figure 5. Core power drift in multi-failure scenario

One of the EDF/CEA blind tests represented the small leakage in the pressurizer, which was the scenario completely absent in the training dataset. The aim was to test the reliability assessment capability of the model. As expected, the model was unable to recognize the faults in this test, but reliability level was correctly set to *low* at the beginning of the unknown situation, which gave the warning to take the model output with care or discard it.

Model was capable to recognize 5 multiple faulty signals simultaneously of all 14 ones. When 6 multiple faults were simulated the reliability level was set to *low* as indication of unreliable output.

Below in Table 1, the model accuracy is illustrated on few selected signals and second error band values for the rated power:

| Signal | rated | Error band |
|-----------------------|-----------|------------|
| Reactor power | 100 % | 0.24 % |
| Temp. control rods | 232 steps | 0.3 % |
| Coolant temperature | 306 °C | 0.1 °C |
| Pressurizer pressure | 155 bar | 0.1 bar |
| Pressurizer level | 64.4 % | 0.3 % |
| Feedwater flow | 2088 kg/s | 5 kg/s |
| Steam generator level | 45 % | 0.012 % |
| Steam pressure | 69 bar | 0.12 bar |

Table 1: Model accuracy at 100 % power (3 simultaneous faults)

6 PEANO as a data validation toolbox

The Neuro-Fuzzy model described in this paper has been implemented in software under Windows NT, in a client/server architecture, as shown in Fig. 6. The system, called PEANO, has the following features:

- PEANO Server:
 1. Full automated training capability. The algorithms described above can be executed and monitored through a friendly user interface, see Fig. 8
 2. Database management. All the training and monitor data can be saved and retrieved in a SQL database, through an ODBC channel.
 3. The server can be connected to the process using one of the following methods:
 - TCP/IP
 - Analog Boards
 - RS-232C
 - From file, for testing
 4. Wavelet based denoising filter of training data

- PEANO client:

1. Up to 20 clients can be connected to the same server, for process monitoring.
2. The monitor display shows instrument values, estimated values, mismatches and reliability levels, both in numeric and trend format (see Fig. 7).
3. Real-time digital filtering, to avoid unnecessary alarms due to noise spikes.
4. Real-time accuracy bands calculation, to provide reliable mismatch warnings.
5. Noise level monitoring.

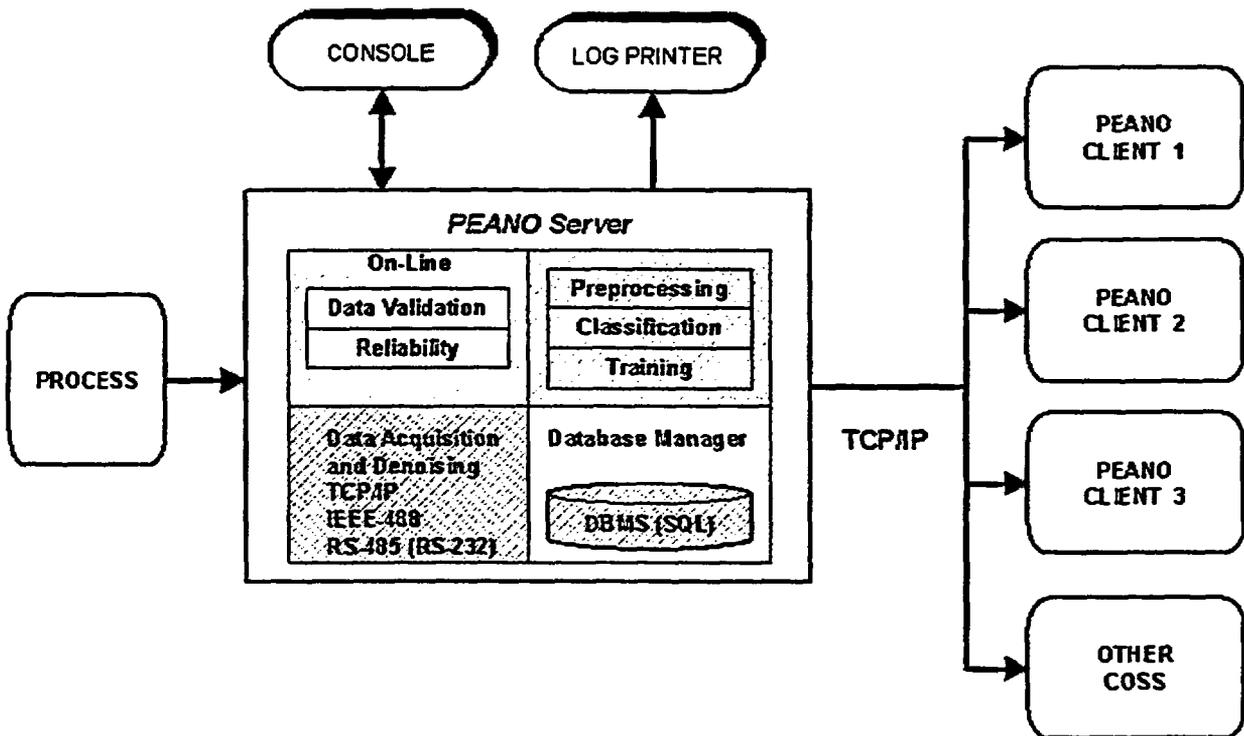


Figure 6 PEANO Client/Server architecture

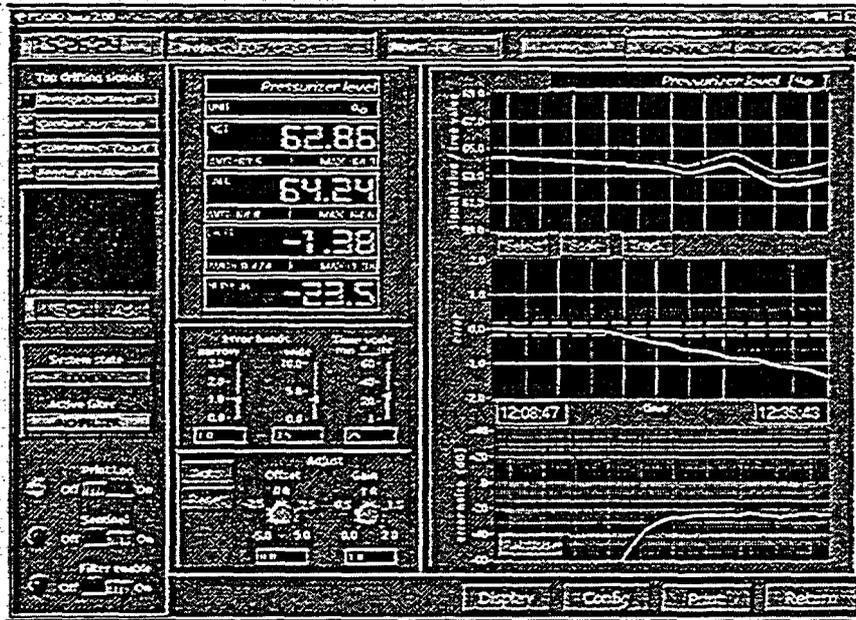


Figure 7 PEANO Monitor display

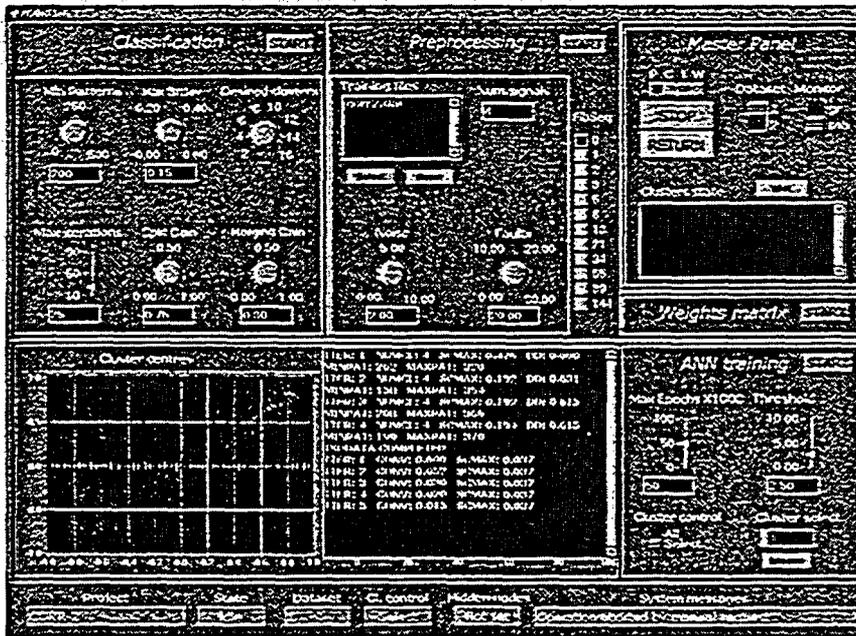


Figure 8 PEANO Training display

7 The ALADDIN extension

ALADDIN is a prototype system that uses a combination of fuzzy clustering and artificial neural networks (ANNs) to approach the problem of classifying events in dynamic processes. The main motivation for the development of such a system derived from the need of finding new principled methods to perform alarm structuring/suppression in a nuclear power plant (NPP) alarm system. One such method consists in basing the alarm structuring/suppression on a fast recognition of the event generating the alarms, so that a subset of alarms sufficient to efficiently handle the current fault can be selected for the operator, minimizing the operator's workload in a potentially stressful situation. The scope of application of a system like ALADDIN goes however beyond alarm handling, to include diagnostic tasks in general. The possible application of the system to domains other than NPPs was also taken into special consideration during the design phase.

In this paper we report on the first phase of the ALADDIN project which consisted mainly in a comparative study of a series of ANN-based approaches to event classification, and on the proposal of a first system prototype which is to undergo further tests and, eventually, be integrated in existing alarm, diagnosis, and accident management systems such as CASH⁵, IDS⁶, and CAMS⁷.

7.1 Related Work

In recent years the range of experimental applications of ANNs to nuclear power plants (NPPs) has been growing steadily and with encouraging results. In particular signal validation has been at the center of ongoing research here in Halden^{8,9}. It is on this body of work and accumulated experience that the ALADDIN project was based.

7.1.1 NPP Transient Classification

Focusing more specifically on transient classification in NPPs, we note that among the first to demonstrate the feasibility of using ANNs were Bartlett and Uhrig¹⁰. That work was developed further and enhanced introducing a modular ANN architecture¹¹. An important contribution was made by Bartal, Lin, and Uhrig¹², where they recognized the necessity for a classifier of being able to provide a "don't-know" answer when presented with a transient of a kind not contained in its accumulated knowledge base.

An alternative way of dealing with temporal data using an *implicit time measure* was proposed by Jeong, Furuta, and Kondo¹³. The same authors recently proposed¹⁴ the *adaptive template matching algorithm* which allows to describe transients in a two-dimensional continuum of time and severity level.

Since the motivation behind ALADDIN derived from the problem of alarm structuring/suppression in an NPP alarm system, we need to mention here the work of Ohga et.al.¹⁵ which integrated a simple ANN based event identification module into an alarm handling system.

In the following we will present a series of neural models for event classification in dynamic processes, with a special emphasis on fast transients and on the ability to produce a "don't-know" classification whenever presented with a previously unseen transient.

7.2 Neural Models for Transient Classification in ALADDIN

7.2.1 Fuzzy Clustering and RBF Neural Classifier

The first approach that was taken consists in a two step process involving a fuzzy and possibilistic fuzzy clustering phase followed by a classification phase. The objective of the clustering is to transform an event description so as to make the classification at the same time simpler and robust (both to noise and to changing initial conditions).

In this context, fuzzy clustering is used to partition the N-dimensional space of the observed variables into regions through which event trajectories pass. The idea is that a record of the regions through which a new event passes could provide enough information to properly classify the event, and still be robust to noise or changing initial conditions. A clustering algorithm derived from *Fuzzy C-Means*¹⁶ was used for this purpose.

One drawback of fuzzy clustering is that even to points which are relatively far from the cluster centers (i.e. far from the regions 'populated' by the trajectories of the events under consideration) are assigned membership values which have to add to 1. This implies that such a clustering will not be able to produce a "don't know" answer in case it is presented with an input vector very different from those it has been designed to recognize, and will assign a high membership to the cluster to which the vector is closest. To overcome this problem a *possibilistic*¹⁷ fuzzy clustering algorithm was developed. Simply stated, *possibilistic* fuzzy clustering relaxes the constraint that the sum of the membership values be always equal to one. The result is that the memberships of points falling in areas not 'covered' by the clusters will all be zero. This property is particularly important if the clustering has to be used for the classification of events of safety-critical systems like NPPs, for which the cost of a misclassification can be very high.

The algorithm which generates the possibilistic fuzzy clustering tends to move the clusters towards the most populated regions of the input space. Relatively to an event classifier (which generates a classification from the memberships), if on one hand this property increases the confidence in the classification (i.e. it reduces the chances of misclassifying an "unknown" event), on the other it can deteriorate its accuracy (i.e. it increases the chances of misclassifying a "known" event). A combination of a fuzzy and a possibilistic fuzzy clustering was therefore chosen, where a fuzzy clustering generates a membership signature for the event classifier, while a possibilistic fuzzy clustering generates a 'confidence signature', i.e. an estimate of how 'common' each vector is for the set of 'known' events.

The actual event classifier is based on radial basis function (RBF) ANNs which, because of their use of local receptive field neurons, are less prone to misclassify 'unknown' cases when compared with the more common perceptron-based ANNs. One problem of both kinds of ANNs is that they are not able to directly model time-dependent data, i.e. they can only map a static input vector to an output value. Our problem is that we want to classify time-dependent fuzzy signatures. The solution adopted here is that of sampling the input data by averaging the fuzzy memberships in three time windows and using those values as inputs to three independent RBF classifiers. Each RBF network will then have one input for each cluster and one output for each event class. The RBFs are trained to recognise a prototypical event for each event class. During operation, the classifications generated by the RBFs are weighted according to a confidence value which corresponds to the maximum possibilistic membership in the respective time windows, and combined to produce a final classification of the event.

7.2.2 Fuzzy Clustering and Cascade-RBF Neural Classifier

In an effort to model more closely the time dependence of the transients we designed a new architecture, still based on RBF ANNs and fuzzy clustering, but with the important difference that the RBFs are not independent of each other but each one bases its classification on the fuzzy memberships in its window (as before) and on the classification generated by the RBF connected to the previous time window. The RBFs become cascade-connected so that the classification of a network gets directly influenced by the classification given by the previous network, and indirectly influenced by all the previous classifications.

The first RBF (i.e. the one connected to the first time window) will be the same as in the previous model, thus having one input for each cluster and one output for each class. Each subsequent network will have one input for each cluster plus one input for each, and one output for each class. As in the previous model, each classification is weighted by the corresponding possibilistic value, while the combination of the classifications is automatically achieved by the connection in cascade.

This architecture should be better able to 'follow' a transient in time as it traverse the fuzzy membership landscape, and possibly eliminate ambiguities which the previous RBF model could not solve.

7.2.3 SOM Neural Classifier

A different approach was devised which is based on the Self Organizing Map (SOM) ANNs¹⁸. The network training iteratively adjusts the neurons weight vectors so as to generate a topographic map of the input space in which similar input vectors activate nearby units. As in the case of the RBF networks, also SOMs are not directly designed to model time-dependent data. Furthermore, SOMs are not classifiers per se so that their output has to be interpreted in order to come to a classification.

The problem of classifying the transients' trajectories becomes here the problem of classifying a sequence of activation patterns in the 2-dimensional map space. To solve this problem we designed a system in which the activation sequences of prototypical transients are stored. During operation, all stored activation sequences are dynamically compared with the activation sequence generated by the current transient, while this evolves in time, to eventually form a classification

7.2.4 Elman Recurrent Neural Classifier

A substantially different approach was attempted which makes use of a special kind of recurrent ANN: the Elman network¹⁹. Recurrent ANNs are a class of ANNs which are able to deal with temporal input signals thanks to an internal architecture which is recurrent, i.e. it makes use of feedback connections among the neurons. This property allows us to feed directly the N-dimensional event trajectories to the ANN and train it to produce in output the required classification.

7.3 Comparative Results

For a comparative evaluation of these models a NPP simulator was used to generate data relative to a small set of events. The simulated plant was the Forsmark 2 BWR, which is a 969-MWe ABB reactor in Sweden, in it was simulated using the APROS simulation environment, developed by IVO and VTT, Finland. The simulated events were: turbine trip with bypass valve operational (TTWBP), turbine trip without bypass (TTWOBP), main steam isolation valve closure (MSIV), feedwater heating loss (FWH), and feedwater controller failure (FWC). Process variables were sampled at 8Hz and the following five were recorded: reactor water level (RWL), feedwater flow (FWF), steam flow (STF), core pressure (CP), and power (P). All the events were simulated at 100%, 80%, 62%, and 49% power. Given only these variables, it is clear that the transients relative to the TTWOBP and MSIV events will be the most similar. In order to study the robustness to noise of the various models, a series of tests were performed by generating noisy transients from the original obtained from the simulator by adding gaussian noise ranging from 5% to 15%. In total 510 transients were in this way generated.

The first test was performed on the RBF model that was trained to recognise the events at 100% power. The general performance was hampered by the anticipated ambiguity between the TTWOBP and MSIV classes. The classification of the FWH and FWC events did not present any problem, while the classification of the TTWBP events degraded gradually with initial conditions getting further and further away from the prototype case. In this model, using also the events at 49% power as prototypes did not improve the performance. The overlap between TTWOBP and MSIV increased and the classification of the TTWBP test cases (i.e. the 80% and 62% cases) did not improve.

The cascade-RBF model general performance was quite similar to that of the simpler RBF model. The more complex cascade-RBF architecture¹ was still unable to properly discriminate between the TTWOBP and MSIV classes, however the classification was sharper, i.e. the separation among the classes was more marked².

With the SOM model, the general performance was similar to the Cascade-RBF model. The SOM architecture³ was better at discriminating between the ambiguous TTWOBP and MSIV classes (at the 100% and 80% power levels), but was less good at generalizing from the prototypes. This led for example to a gross misclassification of the TTWBP events at 62% and 49% power.

The Elman classifier⁴ was the only model which could satisfactorily discriminate between the ambiguous TTWOBP and MSIV classes, and the classification was performed within the first 2.5 seconds, making it a good candidate for alarm filtering applications. The good results of the Elman model are only partly due to the ease with which the Elman ANN is able to use for training several prototypes

¹ In these tests it was based on 10 contiguous time windows and the respective 10 RBFs.

² This comes naturally from the fact that whether in the RBF case two events had to match in only 3 windows (and independently of each order) to receive a similar classification, in the cascade-RBF case two events have to match throughout the development of the transient, i.e. in all the 10 windows and in the correct order.

³ In this case a 6x8 matrix of neurons which was trained on the 100% power level cases.

⁴ In this case a network of eight recurrent neurons which was trained to classify directly the sequences of 5-dimensional vectors derived from the first 20 samples of the simulated events at 100% and 49% plus about other 50 transients with added noise (from 3.5% to 10.5%).

for each class⁵. The main advantage of this recurrent model still lies in its being designed to deal directly with sequences of input vectors.

7.4 The Validation Module and the ALADDIN Prototype

A major problem with most classifiers based on generalization from examples is that their response is not always predictable when they are presented with a previously unseen type of input. The occurrence of an event not included in the design can lead to an incorrect classification⁶. To solve this problem we therefore need a validation module able to confirm that the answer given by one classifier is a meaningful one and not simply coincidental.

The solution proposed here makes again use of possibilistic fuzzy clustering. Here the clustering is made in the space of the whole trajectories, not in the space of the single input vectors. If a trajectory is composed of say m time steps, then the clustering will need to be performed in $m*N$ dimensions. Given such a clustering, a new transient will receive a high possibilistic value only in the case that it is close to the ones used in the training phase⁷.

The tests performed showed results which make this approach a very good candidate for inclusion in a dynamic event classification system, coupled with an accurate classifier such as the Elman model. This is the configuration that we are currently proposing in the ALADDIN prototype, which is schematically shown in Figure 9.

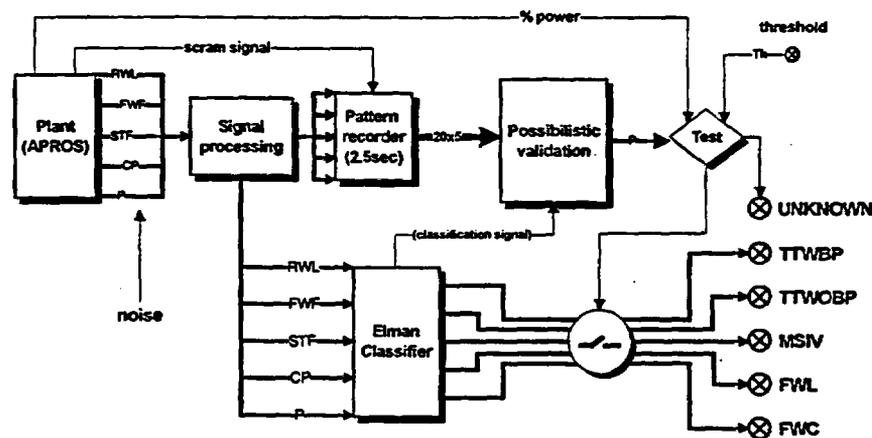


Figure 9: The architecture of the ALADDIN prototype.

⁵ This did not help in the RBF and cascade-RBF cases and was not possible in the SOM case.

⁶ This is true for all the ANNs described, even though at different degrees of seriousness.

⁷ One problem with this validation method is that the $m*N$ clustering space has to be fixed in advance, i.e. all the events need to be synchronized by a trigger signal in our test cases all the events generated a scram signal which was used as a trigger.

In this architecture the decision on whether the classification output of the Elman network is accepted, or an UNKNOWN classification is to be produced, depends on the possibilistic membership value P produced by the validation module, and on whether the plant operating power is within the design range of the classifier.

References

- ¹ Fantoni P F, Neuro-Fuzzy models applied to full range signal validation in NPP, Nucl. Plant Instrumentation, Control and Human-Machine Interface Technologies, NPIC&HMIT'96, The Pennsylvania State Univ., USA.
- ² Fantoni P F, Fignedy S, Racz A, *PEANO, A toolbox for real-time process signal validation and estimation*, OECD Halden Reactor Project report HWR-515, Febr. 1998.
- ³ Gustafson D.E., Kessel W.C., "Fuzzy Clustering with a fuzzy covariance matrix", Proc. IEEE CDC, San Diego, CA, Jan 10-12 1979
- ⁴ P.F. Fantoni, A. Mazzola, "Accuracy Estimate of Artificial Neural Networks Based Models for Industrial Applications", Proc. AI Petro, Lillehammer, Norway, 13-15 Sept. 1995
- ⁵ B.R. Moum, C. Decurnex, N.T. Førdestrømmen, T.V. Karlsen, and T.V. Olsen, "CASH: The New Alarm System in HAMMLAB", *Halden Work Report HWR-480*, May 1996.
- ⁶ T.S. Brendeford and S. Nilsen, "A Sharable Knowledge Repository for Plant Diagnosis Systems", *Halden Work Report HWR-510*, February 1998.
- ⁷ P. Fantoni, Y. Iguchi, G. Meyer, A. Sørenssen, and C. Van Dyck, "CAMS Prototype Extension: Integration of Data Acquisition, Signal Validation, Tracking Simulator, Predictive Simulator, State Identification and Probabilistic Safety", *Halden Work Report HWR-440*, April 1996.
- ⁸ P.F. Fantoni and A. Mazzola, "Multiple-Failure Signal Validation in Nuclear Power Plants using Artificial Neural Networks", *Nuclear Technology*, March 1996.
- ⁹ P.F. Fantoni, "Neuro-Fuzzy Models Applied to Full Range Signal Validation in Nuclear Power Plants", *ANS International Topical Meeting on Nuclear Power Plant Instrumentation, control and Human Machine Interface*, The Pennsylvania State University, 1996.
- ¹⁰ E.B. Bartlett and R.E. Uhrig, "Nuclear Power Plant Status Diagnostics Using an Artificial Neural Network", *Nuclear Technology*, Vol. 97, 1992, pp. 272-281.
- ¹¹ A. Basu and E.B. Bartlett, "Detecting Faults in a Nuclear Power Plant by Using a Dynamic Node Architecture Artificial Neural Network", *Nuclear Science and Engineering*, Vol. 116, 1995.
- ¹² Y. Bartal, J. Lin, and R.E. Uhrig, "Nuclear Power Plant Transient Diagnostics Using Artificial Neural Networks that Allow "Don't-Know" Classifications", *Nuclear Technology*, Vol. 110, 1995.
- ¹³ E. Jeong, K. Furuta, S. Kondo, "Identification of Transient in Nuclear Power Plant Using Neural Network with Implicit Time Measure", *Proceedings of the International Topical Meeting on Computer-Based Human Support Systems: Technology, Methods, and Future*, The American Nuclear Society Inc., La Grange Park, IL, 1995, pp. 467-474.
- ¹⁴ E. Jeong, K. Furuta, S. Kondo, "Identification of Transient in Nuclear Power Plant Using Adaptive Template Matching with Neural Network", *Proceedings of the International Topical Meeting on Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technologies, NPIC&HMIT'96*, The American Nuclear Society Inc., La Grange Park, IL, 1996, pp. 243-250.
- ¹⁵ Y. Ohga, S. Arita, T. Fukuzaki, N. Takinawa, Y. Takano, S. Shiratory, and T. Wada, "Evaluation Test of Event Identification Method Using Neural Network at Kashiwazaki Kariwa Nuclear Power Station Unit No.4", *Journal of Nuclear Science and Technology*, Vol. 33, No. 5, 1996, pp. 439-447.
- ¹⁶ J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, 1981.
- ¹⁷ R. Krishnapuram and J. Keller, "A possibilistic approach to clustering", *IEEE Transactions on Fuzzy Systems*, Vol. 1, No. 2, 1993.
- ¹⁸ T. Kohonen, *Self-Organization and Associative Memory*, 3rd edn., Springer-Verlag, Berlin, 1989.
- ¹⁹ J.L. Elman, "Finding structure in time", *Cognitive Science*, Vol. 14, pp. 179-211, 1990.

Discussion of Comments from a Peer Review of A Technique for Human Event Analysis (ATHEANA)¹

**John A. Forester, Sandia National Laboratories
Ann Ramey-Smith, US Nuclear Regulatory Commission
Dennis C. Bley, Buttonwood Consulting, Inc.
Alan M. Kolaczowski and Susan E. Cooper, Science Applications International Corp.
John Wreathall, John Wreathall & Co.**

Abstract

In May of 1998, a technical basis and implementation guidelines document for A Technique for Human Event Analysis (ATHEANA) was issued as a draft report for public comment (NUREG-1624 [Ref. 1]). In conjunction with the release of draft NUREG-1624, a peer review of the new human reliability analysis (HRA) method, its documentation, and the results of an initial test of the method was held over a two-day period in June 1998 in Seattle, Washington. Four internationally known and respected experts in HRA or probabilistic risk assessment were selected to serve as the peer reviewers. In addition, approximately 20 other individuals with an interest in HRA and ATHEANA also attended the peer and were invited to provide comments. The peer review team was asked to comment on any aspect of the method or the report in which improvements could be made and to discuss its strengths and weaknesses. They were asked to focus on two major aspects: 1) Are the basic premises of ATHEANA on solid ground and is the conceptual basis adequate? 2) Is the ATHEANA implementation process adequate given the description of the intended users in the documentation? The four peer reviewers asked questions and provided oral comments during the peer review meeting and provided written comments approximately two weeks after the completion of the meeting. This paper discusses their major comments.

Introduction

In May 1998, a technical basis and implementation guidelines document for A Technique for Human Event Analysis (ATHEANA) was issued as a draft report for public comment (NUREG-1624 [Ref. 1]). In conjunction with the release of draft NUREG-1624, a peer review of the new human reliability analysis (HRA) method, its documentation, and the results of an initial test of the method was held over a two-day period in June 1998 in Seattle, Washington. Four internationally known and respected experts in HRA served as the peer reviewers. A brief description of the reviewers and their credentials follows:

¹This work was supported by the U.S. Nuclear Regulatory Commission and was performed at Sandia National Laboratories. Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the U.S. Department of Energy under Contract DE-AC04-94AL85000.

- **Dr. Eric Hollnagel** - An internationally recognized specialist in the fields of human reliability analysis, cognitive ergonomics, cognitive systems engineering, and the design and evaluation of man-machine systems. Dr. Hollnagel is the author of more than 230 publications, including six books, articles from recognized journals, conference papers, and reports. In January 1998, he published a book entitled *Cognitive Reliability and Error Analysis Method (CREAM)*, which is itself a new HRA method. He is a member of the Swedish Reactor Safety Council and president of the European Association of Cognitive Ergonomics. Since 1995 Dr. Hollnagel has been principal advisor at the Organization for Economic Cooperation and Development (OECD) Halden Reactor Project, and since 1997 adjunct professor of Human-Machine Interaction at Linköping University, Sweden. He has a Ph.D. in cognitive psychology from the University of Aarhus, Denmark.
- **Dr. Pietro Carlo Cacciabue** - A sector head at the European Commission, Joint Research Centre, Institute for Systems, Informatics, and Safety, in Ispra, Italy. He has published more than 100 papers in professional journals and conferences and is the editor of a number of conference proceedings and books on safety assessment and human factors. Dr. Cacciabue serves as liaison for and holds a number of positions in several international organizations, such as: the International Association for Probabilistic Safety Assessment and Management (director since 1993), consultant for the Direction Générale Aviation Civile, France (since 1994), Institution of Nuclear Engineers, UK, (member since 1984), European Safety Reliability and Data Assoc. (executive committee member 1992-1995), and the European Association of Aviation Psychology (member from 1996 to the present). He has a Ph.D. in nuclear engineering from Politecnico di Milano, Milan, Italy.
- **Dr. Oliver Straeter** - A researcher for Gesellschaft für Anlagen und Reaktorsicherheit (GRS) in Germany in the Safety Analysis and Operational Experience Branch. He was a researcher at the RWTH in Aachen and the Ruhruniversität in Bochum and also worked at Siemens Nixdorf AG compiler laboratory in Munich. Dr. Straeter has published several journal articles in the area of human reliability, including a recent article in *Reliability Engineering and System Safety* (Vol 58, 1997), entitled "Human-Centered Modeling in Human Reliability Analysis: Some Trends Based on Case Studies." Dr. Straeter holds a Ph.D. in human engineering psychology from Technical University of Munich.
- **Mr. Stuart R. Lewis** - A consultant specializing in the application of reliability and quantitative risk analysis methods. Mr. Lewis is the president of Safety and Reliability Optimization Services (SAROS), Inc., Knoxville, TN, which he co-founded in 1984. Examples of current and past relevant work include assisting nuclear licensees in updating and maintaining their probabilistic safety assessments (PSAs) and updating the HRAs for the PSAs of several licensees. He has also assisted the Oak Ridge National Laboratory by reviewing analyses performed under its Accident Sequence Precursor Program, and is assisting Electricité de France in keeping abreast of technical and regulatory developments concerning severe accidents. He performed the HRA portion of several of the probabilistic risk assessments (PRAs) performed by nuclear power plant licensees for the U.S. Nuclear Regulatory Commission's Individual Plant Examination program. Mr. Lewis holds both B.S. and M.S. degrees in nuclear engineering from Purdue University.

In addition, approximately 20 other individuals with an interest in HRA and ATHEANA also attended the peer review meeting and were invited to provide comments. The peer review team was asked to comment on any aspect of the method or the report in which improvements could be made and to discuss its strengths and weaknesses. They were asked to focus on two major aspects:

- (1) The soundness of the philosophy underlying ATHEANA. Are the basic premises on solid ground and is the conceptual basis adequate?
- (2) Is the ATHEANA implementation process adequate, given the description of the intended users in the documentation? Assuming the technical basis is adequate, is the guidance for conducting the search and quantification processes and for integrating the results into the PRA adequate, for example, clear, effective, usable?

The four peer reviewers asked questions and commented orally during the peer review meeting. They also provided written comments approximately two weeks after the meeting. All of the reviewers indicated that the ATHEANA method had made significant contributions to the field of PRA/HRA, in particular by addressing the most important open questions and issues in HRA, by attempting to develop an integrated approach and by developing a framework capable of identifying types of unsafe actions that generally have not been considered using existing methods. The reviewers had many (and sometimes similar) concerns about specific aspects of the methodology and made many recommendations on ways to improve and extend the method and to make its application more cost effective and useful to PRA in general.

This paper discusses the major comments received from the peer review team and provides responses (but not necessarily resolutions) to specific criticisms and suggestions for improvements. A list of the general strengths and weaknesses of ATHEANA, as indicated by the reviewers, is provided first. Next, specific comments bearing on major aspects of the method are presented and discussed. Finally, general comments related to improving the efficiency and usefulness of ATHEANA are addressed.

General Strengths and Weaknesses of ATHEANA

The reviewers' general opinion of ATHEANA is that the method represents a significant improvement in HRA methodology; it is a useful and usable method; and it is a "good alternative to first-generation HRA approaches." However, the method does not yet go far enough and therefore needs to be improved and extended. Several of ATHEANA's strengths, as indicated by the four reviewers, are listed below.

- 1) "Until now, in my opinion, there is no other published approach that tries to solve the problem of including EOC [errors of commission] in PSA in such an extensive way. Other methods address only parts of this. Overall, the general approaches and concepts developed in the ATHEANA-method are appropriate to deal with the problem of EOC. I think that the ATHEANA-method as currently documented contains a lot of important aspects for understanding and integrating EOCs into PRA. However, many aspects are only mentioned implicitly. An explicit and concise elaboration is necessary to assure practicability..."
- 2) "The real value of ATHEANA seems to be as a systematic way of exploring how action failures can occur. This is something that conventional HRA methods do not do well, if they do it at all, since they tend to focus on producing numbers. Although this use of ATHEANA does not really answer the need for an HRA approach, it might have a value in itself (as the comments from the demonstration participants expressed) and it might conceivably be decoupled from the HRA side. In that case a more streamlined method may be developed, that is less cumbersome to use. The demonstration of ATHEANA very clearly showed how it can be used to develop detailed qualitative insights into conditions that may cause problems, how it may generate a solid basis for

redesign of working procedures, training, and interface, and how it may be used as a tool for scenario generation. Each of these are significant achievements in their own right.”

- 3) **“The method described in ATHEANA is certainly well suited for overcoming the difficulties encountered when applying more classical human reliability methods and focuses on the important issues of context and cognition that need to be tackled. Many aspects of the methodology are commendable and give great added value to the whole methodology. In particular, the following features are important:**
- **the details in describing many processes and steps in the application of the methodology;**
 - **the consideration for the crucial features that affect human cognition and behaviour in managing modern plants, included in concepts like the error-forcing context; and**
 - **the identification of the appropriate retrospective approach for the evaluation of the factors influencing behaviour and basic data for prospectively analysing the likely outcome of erroneous behaviour and probabilities.”**
- 4) **“Properly applied, the methods that comprise ATHEANA should be able to yield significantly more insight into the nature of human actions that can contribute to the occurrence of a core-damage accident. These methods clearly provide a framework for identifying some types of unsafe actions, and especially errors of intention, that would generally not have been considered using current methods. Moreover, they allow for a much more careful definition of the context and causes of these unsafe actions.**

Without broader application of the methods, however, it is impossible to draw conclusions regarding the degree to which important actions that are not considered in present PRAs will be identified. It is reasonable to expect that some of the most important potential unsafe actions would be the result of subtle aspects relating to interactions among plant conditions or performance shaping factors that would be very difficult to postulate, even with the proper team makeup and extensive time available for the analysis.

What can be expected is that the methods will provide for the integration of understanding from the diverse team members that will lead to these new insights. This should be a synergistic process, allowing knowledge to be shared and captured in a way that enhances both the completeness and realism of the PRA, and the quality of training and procedures. A significant advantage of the method could be to provide a rationale for the characterization of the human failure events that often eludes us in present PRAs. While present methods may arguably yield reasonable quantitative results, they often fail to provide an understanding of the underlying causes of the human failures that are analyzed. Absent that understanding, it is very difficult to identify measures that can be taken to reduce the risk associated with unsafe actions. Consequently, it is often frustrating to identify a human action as risk-significant, but not to be able to give very satisfactory answers as to why, or what could be done to reduce that significance. With ATHEANA, on the other hand, the analysis of an unsafe action is necessarily truncated if an error forcing context cannot be identified.”

The above statements clearly indicate that the ATHEANA method has made significant improvements in HRA methodology and that the method, as documented, is a useful and usable tool. Perhaps not surprisingly, current members of the ATHEANA development team (the authors of this paper) agree generally with the above statements. However, the reviewers were also very clear in indicating that, in their opinion, there are several important general shortcomings of ATHEANA. These are listed below.

- 1) "There seems to be an inconsistency in the level of models being used, ranging from EOO-EOC (errors of omission - errors of commission) over the information processing model to the notion of slips and mistakes. It would be interesting to consider how the search process could be strengthened while relaxing the dependence on the model(s)."
- 2) "There is no identifiable way of encompassing management and organization [M&O] factors or responding to the challenges of the broader socio-technical or contextual way of thinking (which also is seen by the conceptual problems in taking M&O factors into account in PSA)."
- 3) "Insufficient consistency in the terms and concepts used, and significant differences between what is written in NUREG-1624 and what was said at the review."
- 4) "The ATHEANA method is very cumbersome and presumably very costly. The guidance is too complex and depends too much on subject matter experts."
- 5) "The quantification method is weak, and the quantitative results (of the demonstration) are unsubstantiated. The quantification is excessively dependent on expert judgement, hence possibly has low reliability as a method."
- 6) "The qualitative results are good, but these might have been obtained in other ways, perhaps more efficiently. It is also doubtful whether a utility will undertake a significant effort just to get the qualitative results."
- 7) "The implementation of the basic approaches is sometimes not elaborated far enough from my perspective. This makes the use of the method in the current status difficult and may cause high variance between different users. I also observed that the document NUREG-1624 and the presentations on the peer-review are sometimes not in accordance to each other. In order to have a usable and profound method, the basics has to be refined and extended."
- 8) "Especially, I see the danger that the whole suggested procedure may fail if the role of the cognitive model (i.e. to work out and structure EMs [error mechanisms]) is not elaborated further. The cognitive model has a considerable effect on the consistency between EMs, the compatibility of prospective and retrospective analysis, the link between EFC [error-forcing context], EM and UA [unsafe actions] as well as the quantification procedure."
- 9) "The methodology clearly presents a dilemma. Its effectiveness results from forming a diverse, experienced project team to perform a comprehensive, broad-ranging analysis. Few organizations, however, appear to be in a position to undertake such an extensive analysis without clearly defined, commensurate benefits. Thus, even if it is an excellent methodology from a technical standpoint, it will not be very valuable if it will not be used."

- 10) "The potential wide application and popularity of the method are, however, associated with the *easiness of application* of the method and the *completeness* of the supporting information and data. The first issue (*easiness of application*) is related to the clear differentiation between retrospective and prospective analysis, which contains also the question of applicability of the cognitive model. The method, as presented in the report, generates some confusion, especially for non-specialists in human factors, even though one could argue that the ATHEANA team should contain such expertise. The question of the availability and *completeness* of a reference database and clear tables of parameters and variables sustaining the HRA approach has, in practice, already been almost completely tackled and solved. What remains to be done is simply the clear definition of the connections between such databases and parameters on the one hand and models, paradigms and structure of ATHEANA on the other."

Although the above set of comments is not necessarily complete in regard to the limitations of ATHEANA as indicated by the peer reviewers, it is thought that the selected set does represent the more important general limitations identified by the reviewers. Some of the above criticisms are responded to directly, but in other cases, some future decisions are required. The criticisms and responses are grouped below according to major aspects of ATHEANA.

The ATHEANA Framework and Underlying Models

Two important aspects of the ATHEANA methodology are (1) the multi-disciplinary HRA framework (see Figure 2.1, NUREG-1624 [Ref. 1]) that describes the interrelationships between human error mechanisms, the plant conditions and performance-shaping factors (PSFs) that set them up, and the consequences of the error mechanisms in terms of how the plant can be rendered less safe, that is, UAs and (2) the human information processing or "cognitive" model (see Figure 4.1, NUREG-1624 [Ref. 1]) that is used to describe the human activities and mechanisms involved in responding to abnormal or emergency conditions and thereby assist analysts in searching for potential unsafe human actions. Several of the criticisms listed above (e.g., 1, 8 and 10) raise concerns about the descriptions and use of the framework and the cognitive model in ATHEANA. Essentially all of the peer review team had questions or concerns about these aspects of ATHEANA.

Regarding the multi-disciplinary HRA framework, several reviewers thought that the definitions and distinctions between the components of the framework and their interrelationships with each other and with the cognitive model were not sufficiently clarified. The reviewers considered this important because they correctly assumed that understanding the framework (and to some extent its relationship with the cognitive model) was important to understanding the ATHEANA methodological approach. One concern was exactly what was meant by "error mechanisms," how they are used in ATHEANA, and whether or not the terminology was appropriate, given the underlying assumptions of ATHEANA, for example, people usually behave rationally and are led to UAs as a function of the circumstances. Another concern was that the distinction between error mechanisms, PSFs and plant conditions was not sharp enough.

Clearly, "crisper" definitions of these terms are needed in the ATHEANA documentation because they are used to guide analysts in their search for UAs and the associated EFCs. One goal of using the construct of error mechanisms is to convey to analysts that there are human information processing activities that may be appropriate in some circumstances, but not in others. Examples of such activities

are provided in the ATHEANA documentation and they are elaborated to some degree in the discussion of the cognitive model (Section 4 of NUREG-1624). The main purpose of the discussion in Section 4 is to encourage analysts to think about the potential for human error in a different manner than has been done in other HRA methods and not necessarily to provide a complete and validated set of error mechanisms. It is not obvious that further elaboration of possible error mechanisms will necessarily facilitate the ATHEANA search process or the quantification process. Nevertheless, the clear use of the construct of "error mechanisms" in the context of ATHEANA will be addressed. To the extent that additional explanation and elaboration of potential error mechanisms will facilitate the search and quantification processes, such work will be performed for later revisions.

Consideration will also be given to a couple of reviewers' suggestion that the term "error mechanism" should be dropped because human information processing is probably not limited only by processing "mechanisms," which implies structures, (e.g., processing is probably also limited by inappropriate processing strategies) and because the behavior that leads to UAs is only an "error" in hindsight. As is assumed by ATHEANA, the information processing performed may have been perfectly appropriate in most situations and is inappropriate only because of special circumstances; it therefore is not an error in the usual sense. Recommendations for a replacement term for the construct included "behavior mechanisms" or simply "cognition."

As noted earlier herein, another concern expressed by the reviewers was with the distinction between plant conditions, PSFs, and error mechanisms. It was argued that it is not always easy to determine whether a particular factor belonged in one category or another (e.g., whether procedures and instrumentation problems should be categorized as plant conditions or PSFs) and that it was necessary for ATHEANA to make the distinctions clear. One reviewer indicated that the PSFs should be standardized and made complete. The current ATHEANA documentation has acknowledged that, in some cases, the distinctions are not always perfectly clear, but the emphasis from the analysis point of view is to ensure that the factors relevant to the EFCs are considered. Although it may be possible for the ATHEANA team to develop a useful underlying model for grouping the relevant factors and this effort may be attempted for revisions to the method, the main consideration in the application of ATHEANA is that as many relevant factors as possible are considered in identifying the EFCs.

Other issues regarding the models used in ATHEANA concerned the use of the EOC-EOO distinction, the slips versus mistakes categorization in the context of the other models used in ATHEANA (e.g., see criticism 1), and the ability of the method to correctly consider crew-related factors when the cognitive model generally applies to information processing by an individual. The latter concern suggests that it might be useful to include a "crew interaction" model that could be integrated with the cognitive model. The team will examine the feasibility and usefulness of such an endeavor.

Regarding the slips versus mistakes categorization, several reviewers argued that this categorization was probably not necessary and at least one argued that it was inappropriate. The use of such terminology, which does presume an underlying model not explicitly adopted by ATHEANA, will be addressed in future revisions.

Finally, several reviewers also suggested that the framework and models used in ATHEANA be compared to other more familiar models from existing methods in order to elucidate the differences between ATHEANA and other HRA approaches. This would certainly be a useful addition to the ATHEANA report in that it would assist analysts in realizing the advantages to conducting an

ATHEANA HRA. Clearly, revision of the ATHEANA documentation should discuss the uses and appropriate application of ATHEANA to various analysis tasks.

The ATHEANA Process

This section addresses a variety of important comments on aspects of the ATHEANA process.

Retrospective Analysis

The use of an ATHEANA-driven retrospective analysis of plant and other operational events was listed as one of the strengths of the ATHEANA process (see strength 3). More than one of the reviewers commented on the positive aspects of the use of retrospective analysis for assisting analysts in evaluating their plant and supporting the proactive HRA. In fact, their main concern was that a formalized, structured procedure, separate from the proactive search process detailed in ATHEANA, was not provided in the existing documentation. They suggested that a separate write-up and flow diagram be developed on how to perform retrospective analysis and on how it interfaces with the proactive analysis. Reviewers concerned with the definitions and relationships/connections between the elements in the framework and cognitive model also felt that clarification of these aspects would also greatly facilitate the retrospective analysis (see criticism 8). They argued for "taxonomies for actions, errors, and PSF" and clear rules for event decomposition in the retrospective analysis. In addition, they also suggested providing improved guidance on how to use the HERA database (Ref. 2) and the retrospectively analyzed events documented in Appendix B of NUREG-1624. [Note that HERA is a database being developed for the USNRC that contains documentation of significant events from nuclear and other industries. The events are represented from the ATHEANA perspective and in ATHEANA terminology.]

The ATHEANA team agrees that additional guidance on how to perform and use retrospective analysis and the HERA database would be useful additions to the ATHEANA documentation. Analysts would be able to learn more directly about the characteristics of ATHEANA and in addition to "self-training" on the ATHEANA "philosophy," framework, and models, they would better understand events that have occurred at their plant and how other events might occur in the future.

Prioritization Process

Several of the criticisms listed above (e.g., 4, 6, and 9) indicate that the demands of applying ATHEANA may be cost and time-prohibitive for many nuclear power plants. One aspect of ATHEANA that was developed in an attempt to allow users to focus their limited resources was a process for prioritizing the more important accident scenarios. While the reviewers generally were supportive of the prioritization process, several suggested that the process be further improved and proceduralized. Specifically, they wanted a "greater consideration of the risk potential of possible human failure events (HFEs)" and (on the basis of information provided at the peer review on the results of the trial application of ATHEANA) an earlier identification and assessment of crew characteristics and other M&O factors that might make certain types of scenarios more likely to contain risk significant UAs than others.

Once again, the ATHEANA team agrees that improvements in the prioritization process, as suggested by the reviewers, would be useful. A characterization of the way plant crews interact with one another and approach accident scenarios would assist analysts in determining the types of scenarios likely to be

problematic (see Appendix A, Section A.7, of NUREG-1624 for details). Explicit incorporation of other M&O factors (which is considered a weakness of ATHEANA; see criticism 2) at the prioritization stage may also be beneficial. It should be noted that there is nothing about ATHEANA that is inherently incompatible with the consideration of M&O factors (contrary to criticism 2). The main problems associated with accounting for M&O factors in ATHEANA are that there are no currently accepted methods for modeling such factors, and the costs associated with the additional analysis may offset the benefits.

In addition to these two items, there were several other comments related to the ATHEANA process that the ATHEANA team, in principle, agree with. They include the following:

- Provide further guidance for the creative thinking/search process to lessen variability and interpretation, including providing guidance on how to "manage" group discussions. Also emphasize the need to document the process "as you go" and more closely link the documentation tables with the relevant sections of the search process.
- Stress more strongly the importance of modeling the support systems, in addition to the main safety systems, in searching for potential HFEs and UAs.
- Discuss to what extent dynamic reliability is or is not part of the process and why.
- Further stress where and how one treats organizational factors, team interactions, recovery, and dependencies

One additional comment on the ATHEANA process warrants a response from the ATHEANA team. It was suggested that there should be an explicit use of formal task analysis in conducting ATHEANA. While it is true that some of the existing HRA methods recommend the use of formal task analysis in order to understand the operators' tasks during accident scenarios, it is not clear that the additional costs associated with formal task analysis would necessarily be useful in applying ATHEANA. In conducting ATHEANA, the HRA team, using appropriate procedures, examines the crew's responsibilities during various accident scenarios and, when possible, conducts simulator exercises. It may be beneficial, however, to emphasize the step of carefully examining procedures relevant to particular accident scenarios early in the process of identifying potential UAs and their EFCs. This step is certainly part of task analysis and should assist analysts in identifying the more critical and likely UAs for further analysis.

The ATHEANA Quantification Process

The reviewers raised several issues associated with quantification. These include the overall ATHEANA approach of identifying and quantifying situations where the likelihood of failure is very high, the methods used to quantify a UA in a particular EFC, and the effect of the various PSFs and plant conditions on the likelihood of failure. Other comments pertained to the need to address recovery actions and dependencies in the quantification process.

A basic premise driving the development of ATHEANA is that the HFEs that have heretofore been most problematic for identifying and assessing their impact on plant risk are those in which a particular context creates a very high likelihood of failure. This premise is in contrast to the premise implicit in

most other HRA methods that there is a constant (and usually low) likelihood of human failure for any given accident scenario. (It is true that some HRA methods have moved beyond this simple assumption, but they have not been widely used and have rarely been applied in a systematic way.) Therefore, the search process and the associated quantification process are principally aimed at identifying those conditions in which the UA probability will be much higher than in other non-forcing conditions. However, this fact does not imply that the application of ATHEANA would never identify situations in which the probability of the UA, given the EFC, is significantly less than 1.0. In such situations in which human error probabilities must be estimated, existing applicable HRA methods may be useful for quantifying the error probability, given the defined EFC.

Several reviewers suggested that the methods for estimating the probability of the UA be revised or broadened. We agree that alternative methods can be used. In the trial application, HEART (Ref. 3) was used because it most directly used conditions similar to those identified as EFCs in the scenarios, bearing in mind the data sources used in HEART and the level of description for the conditions under which the data were gathered. It is important to ensure that the method and data used to quantify the likelihood of an unsafe action in a particular EFC will be sensitive to those factors that create the forcing nature of the EFC conditions. An alternative approach that was suggested is to use a subjective-assessment method like SLIM-MAUD (Ref. 4). Such methods could be used in principle. However, the continuing difficulty is one of selecting appropriate anchor points for the assumed probability distribution. This problem has been raised previously in reviews of HRAs that have used methods like SLIM-MAUD in which the analyst provides the range within which a point probability is interpolated.

One reviewer suggested the use of tables for specific PSFs and plant conditions that showed their influence on the likelihood of unsafe actions. Such data could be derived from historical experience in the events reported in the database. However, this approach is at odds with the ATHEANA method, which considers the influence of PSFs and plant conditions to be an integral set of influences on performance, and not separable and discrete influences such as those reported in THERP (Ref. 5). In ATHEANA, the typical issue is "What combination of plant conditions and weaknesses in the displays, procedures, etc., has to occur to mislead operators into believing that action 'x' needs to be taken?" The key is that it is the combination, not each influence separately, that is important.

It is agreed that the analysis of recovery actions is problematic. In applying ATHEANA, the team has considered recovery on a case-by-case basis, looking specifically at ways the scenario may develop, where additional outside staff may become involved, and so on. The approach thus far has not been to treat recovery actions as separate from the initial UAs. Similarly, the method does not include explicit processes to model and quantify dependencies between actions. Clearly, future revisions and applications of ATHEANA must better address the analysis of recovery actions and dependencies.

Improving the Efficiency, Usefulness, and Consistency of ATHEANA

Several of the comments from the reviewers (e.g., criticisms 4, 6, and 7) express concerns about the resources required to apply ATHEANA and whether or not the obtained results will be important enough and complete enough for users to justify the costs. A related concern is whether the method has been specified in enough detail and "elaborated far enough" to allow consistency in the results obtained by different analysts applying the method. Similar concerns regarding resource demands and completeness were raised by the participants of the first demonstration of ATHEANA, which was held in 1997 at a

pressurized water reactor nuclear power plant (see Appendix A, Section A.7, of NUREG-1624 for details).

The ATHEANA team acknowledges that a broad and careful application of ATHEANA will require significant resources. Although the search for important HFEs, UAs, and their EFCs will never be trivial, it can be manageable. Thus, steps will be taken to improve its efficiency (some of which are discussed below). Will the resources demanded by the method be worth it? ATHEANA will identify demanding accident scenarios and potential UAs and EFCs that could lead to serious accidents. Whether or not the method will identify numerous events that result in large increases in calculated plant risk metrics remains to be seen. Moreover, given the inadequacies of the HRA methods that were used to conduct the existing nuclear plant PRAs, it is impossible to know exactly what a realistic estimate of the baseline HRA contribution should be. Therefore, it is difficult to predict what kinds of changes in risk metrics to expect. In any case, the benefits of ATHEANA are much broader than those from performing revised PRA calculations alone. The improvements in HRA modeling to better identify operator vulnerabilities in accident scenarios and to better understand what are the contributors to operator performance will certainly be of significant benefit in assessing and managing plant risk. Nevertheless, it must be the case that the method can be applied without an excessive demand on licensee resources.

The peer reviewers and others identified several actions that will increase the effectiveness and efficiency of ATHEANA. These actions include the following:

- developing a computer-based user support system to guide the process and the documentation of the results,
- refining the prioritization process to facilitate identification of the types of scenarios and situations most likely to create problems,
- developing better guidance on when and how to develop and use simulator exercises to learn as much as possible about where and how unsafe actions can occur, and
- producing a "quick reference guide" that would allow analysts to bypass reliance on the NUREG document once they have some experience with the method.

Another issue raised by the peer reviewers concerns consistency in the application of the process and the potential for significant variability in results because of some of the "open-ended" aspects of ATHEANA, (for example, the creative thinking and brainstorming aspects of the process for identifying EFCs and the use of expert judgment in the quantification process). The ATHEANA team agrees that additional guidance is needed to ensure consistency in the results obtained using the method.

Finally, it should be noted that reviewers of the method suggested that the documentation provide estimates of the costs and resources required to perform ATHEANA and that criteria should be provided for when ATHEANA should be used. While the former suggestion may be difficult to implement until additional tests of ATHEANA are completed, it is a reasonable suggestion. Providing a listing of criteria for when use of ATHEANA is called for would seem to be straightforward and will be considered for the revision.

Other Useful Suggestions

Several other comments received from the peer review team are worth noting because they are good suggestions that would improve ATHEANA. They include the following:

- ATHEANA should include an overview of PRA for participants without a background in PRA. Any training programs developed for ATHEANA could also provide such an overview, and aspects of PRA could be treated in more detail as the analysis progressed.
- It was recommended that a single “running” example be used while discussing the implementation process.
- It was recommended that additional examples for BWRs should be added. PWRs are overemphasized.

Conclusion

Taken together, the comments from the peer review team indicate that the work performed in the development of ATHEANA has resulted in significant contributions to the field of HRA and that ATHEANA is a viable HRA method. However, the reviewers also indicated that there were important clarifications and improvements that needed to be made to ATHEANA. Clearly, many of the recommendations made by the reviewers would, if implemented, make ATHEANA a better, more effective, easier to use, and more “encompassing” methodology. However, a number of factors must be considered in determining which of the suggested changes are necessary, which would be useful but are not critical, and which would be useful but are currently impossible. The development of an HRA method such as ATHEANA is certainly limited by the state of current knowledge in a number of domains such as cognitive psychology, crew dynamics, and management and organizational factors. In addition, the unavailability of actual data from crew performance in nuclear power accidents or from other domains that might be generalized to control room performance certainly limits the ability of any HRA method to precisely predict performance. Other factors include the danger of over-complicating the method in attempts to be more precise and complete. It seems to the ATHEANA team that the most important goal is to provide a usable method that is as cost-effective as possible -- one that will allow analysts to identify, understand as much as possible, and quantify as accurately as possible, potential unsafe human actions that could lead to serious accidents in nuclear power plants or other domains. The explicit procedures, information, and guidance provided in ATHEANA certainly provides HRA analysts with a new and explicit set of tools to achieve this goal. To the extent viable changes recommended by the reviewers will further this goal, in particular by making the method more valid and easier to use, attempts will be made to incorporate them into the ATHEANA methodology.

References

1. U.S. Nuclear Regulatory Commission, *Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis, Draft Report for Comment*, NUREG-1624, Washington, DC, May 1998.
2. S. E. Cooper, W. J. Lucas, and J. Wreathall, *Human-System Event Classification Scheme (HSECS) Database Description*, BNL Technical Report No. L2415/95-1, Brookhaven National Laboratory, Upton, NY, December 1995.

3. J. C. Williams, "A Data-based Method for Assessing and Reducing Human Error to Improve Operational Performance," 1988 IEEE Fourth Conference on Human Factors and Power Plants, Monterey, California, IEEE, 1988.
4. D. E. Embrey, et al., "Slim-Maud: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgment," Vols. 1-2, NUREG/CR-3518, U.S. Nuclear Regulatory Commission, Washington, D.C., March 1984
5. Swain, A.D., and H.E. Guttman, *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*, NUREG/CR-1278, Rev. 1, Sandia National Laboratories, Albuquerque, NM, August 1983.

Incorporating Aging Effects into Probabilistic Risk Assessment

Curtis L. Smith¹, George Apostolakis,² Tsu-Mu Kao,² Vik Shah¹

Abstract

Traditional probabilistic risk assessments (PRAs) of light water reactors (LWRs) include active components but not passive systems, structures, and components (SSCs) because the passive SSCs are much more reliable than the active ones. The objective of the paper is to evaluate the feasibility of incorporating aging effects into PRAs so that risk impact of aging can be estimated. We have evaluated this feasibility by incorporating into PRA a flow-accelerated corrosion model, the KWU model developed by Kastner and Riedle, which estimates wall thinning. For this purpose, we have used the PRA model developed for the Surry Individual Plant Examination. In addition, we have employed a *load-capacity* model based upon a *reliability physics* model and simplifying assumptions for estimating failure caused by flow-accelerated corrosion.

We have demonstrated that the results and insights gained from the U.S. Nuclear Regulatory Commission (NRC) Nuclear Power Aging Research Program and other NRC and industry programs related to materials degradation can be integrated into the existing PRA models. However, our evaluation represents feasibility of such integration and *should not* be construed to represent either relative or absolute magnitude of risk posed by flow-accelerated corrosion in PWRs.

1. Idaho National Engineering and Environmental Laboratory (CLS2@INEL.GOV)
2. Massachusetts Institute of Technology (APOSTOLA@MIT.EDU)

1. Introduction

The high initial reliability of passive systems, structures, and components (SSCs) at operating light water reactors (LWRs) is being reduced as the plants are getting older. Deleterious effects of aging mechanisms have been manifested at older nuclear power plants. While not accounted for in the design of passive SSCs, several aging mechanisms and affecting phenomena have caused SSC damage (Shah and MacDonald 1993). This damage has raised questions about the continued safety and viability of older nuclear power plants. The U.S. Nuclear Regulatory Commission (NRC) sponsored the Nuclear Plant Aging Research (NPAR) program during 1985-1994, which has gathered data and developed insights in aging of LWR SSCs (Vora 1994). Several other NRC- and industry-sponsored projects have also generated both qualitative and quantitative information related to material degradation in passive SSCs. This large collection of information, however, has not been used to estimate the risk impact of aging of passive SSCs.

The work evaluates the feasibility of incorporating this large body of qualitative information into probabilistic risk analysis so that the risk-impact of aging of passive SSCs can be assessed. Simpler approaches, such as the linear failure rate model (Vesely, 1987), are not considered despite ease of their application because they can not utilize the material degradation information collected by NPAR and, therefore, resort to expert opinion for the estimation of their parameters. The incorporation of aging models into PRA will allow assessing the effect of aging on core damage frequency. Such incorporation will also provide stronger technical basis for making decisions related to plant operation and maintenance. For example, the results may be used in risk-informed inspection to prioritize components and optimize inspection activities. In addition, the risk-informed methods presented in the paper will allow better utilization of resources geared toward plant operation.

The overall objective of the work discussed in this paper is to assess the *feasibility* of applying the *LANL/ASCA* method for incorporation of reliability-physics based models, expert judgement, and the results of the Nuclear Plant Aging Research (NPAR) program into an integrated aging risk assessment. The *LANL/ASCA* method refers to the application of PRA methods described in NUREG/CR-6157, *Survey and Evaluation of Aging Risk Assessment Methods and Applications* (Sanzo et. al. 1994). The feasibility assessment that is presented here was performed via a trial application of the NUREG/CR-6157 methodology. This application utilizes an existing nuclear power plant PRA for which a SAPHIRE computer model currently exists [Smith et al., 1998]. The main focus of the project is on pressurized water reactor (PWR) components, but the application presented here can equally be applied to boiling water reactor (BWR) components.

We first summarize the aging mechanisms acting on the major PWR components. We then briefly describe the flow-accelerated corrosion (FAC) mechanism. Next we review the related field experience to identify how FAC of carbon steel secondary piping may impact plant risk. We then present the approach developed to incorporate the FAC model into PRA and apply it to four cases related to FAC-caused failures in PWR secondary piping. We follow this with the demonstration of the approach by applying it to selected segments of feedwater piping using the SAPHIRE software and the Surry Independent Plant Evaluation (IPE) model. Finally we present conclusions.

2. Aging Mechanisms in PWR Plants

The major PWR components, such as reactor pressure vessel, reactor coolant system piping, steam generator tubes, feedwater and main steam lines, are subject to several age-related degradation mechanisms. The main degradation mechanisms may be divided into four categories: embrittlement, fatigue, stress corrosion cracking, and corrosion. Embrittlement mechanisms include radiation embrittlement of reactor pressure vessels and thermal aging of cast stainless steel piping. Fatigue mechanisms include low- and high-cycle thermal and mechanical fatigue of piping and other components. Stress corrosion cracking mechanisms include intergranular and transgranular stress corrosion cracking of, for example, stainless steel components, and primary water stress corrosion cracking of Alloy 600 components. Corrosion mechanisms include, for example, general corrosion, FAC, and boric acid corrosion of carbon steel components. Embrittlement mechanisms reduce material fracture toughness whereas other mechanisms reduce component strength.

Review of service experience reveals that the major components are experiencing aging-related degradation of material properties or reduction in strength. In addition, several degradation mechanisms and loading conditions were not anticipated in the original design but have caused failures in the field (Gosselin 1997, Shah et al. 1998). An example of such failures include a rupture of feedwater piping caused by FAC. For effective aging management, it is essential to evaluate the risk impact of the degradation mechanisms. Physics-based models may be incorporated into the PRA for such evaluation. Toward that end, the main degradation mechanisms are summarized below.

2.1 Radiation Embrittlement. The main component affected by radiation embrittlement is the reactor pressure vessel, which is fabricated from low-alloy ferritic steels. The radiation-induced microstructural changes cause an increase in the yield and ultimate tensile strengths, a decrease in the fracture toughness, an increase in the ductile-to-brittle transition temperature (below which the vessel will behave in brittle manner), and a drop in the upper shelf energy (lower toughness at operating temperature). The increase in the transition temperature of vessel materials depends on chemical composition (especially presence of trace elements, copper, nickel, and phosphorous), neutron fluence ($E > 1$ MeV), temperatures, and post-weld heat treatment. Several computer codes have been developed for evaluation of structural integrity of reactor pressure vessel. Only one of these codes, VISA-II Code (Vessel Integrity Analysis Code), is in public domain. The VISA-II Code performs both deterministic and probabilistic analysis to determine reactor vessel failure probability following through-wall cracks caused by PTS events (Simonen et al. 1986a, b).

2.2 Thermal Aging. Several components in the PWR reactor coolant system are made from duplex austenitic-ferritic stainless steels, also called cast stainless steels: main coolant piping in several Westinghouse-designed PWRs, fittings in larger diameter austenitic stainless steel piping, and reactor coolant pump and valve bodies in most PWRs. The ferrite present in the cast stainless steel improves its resistance to sensitization, stress corrosion cracking, and intergranular corrosion. However, the presence of ferrite also causes upward shift in ductile-to-brittle transition temperatures and reduction in fracture toughness after long-term exposure at PWR operating temperatures. The extent of this degradation is higher with longer exposure time, higher exposure temperature, and higher ferrite content. This degradation mechanism is termed *thermal aging*. Piping made from wrought austenitic stainless steel (for example, Type 304 stainless steel) is not susceptible to this mechanism because this material does not contain ferrite.

Chopra (1992a,b) has developed two different approaches to determine the extent of thermal aging of cast stainless steel. These approaches quantify the extent of aging at the PWR operating temperatures [288°C

(550°F)] by measuring the room-temperature Charpy impact energy after aging at temperatures in the range of 300 to 400°C (570 to 750°F). Chopra (1992b) has also developed a procedure for estimating mechanical properties, i.e., Charpy V-notch energy and elastic-plastic fracture toughness, of thermally aged cast stainless steel piping components.

2.3 Low-Cycle Fatigue. Fatigue damage occurs only in regions that deform plastically under the influence of fluctuating loads. Thus, smooth specimens free of stress risers that are subject to elastic stress fluctuations do not experience fatigue damage. On the other hand, fatigue damage does occur under elastic stress fluctuations if the component contains stress risers where the localized stresses and strains exceed the elastic limit of the material. After a certain number of load fluctuations, the fatigue damage at the regions of stress concentration causes initiation and subsequent propagation of fatigue cracks in the plastically deformed region.

Thermal and mechanical stresses imposed during system transients, including heatup and cooldown, cause low-cycle fatigue damage in the PWR vessel and piping components. The susceptible sites in the PWR primary coolant systems are the nozzles, dissimilar metal welds, and elbows. There have been no reported failures, not even discovered cracks, in existing PWR main coolant piping.

The ASME Code provides fatigue design curves that are entirely based on data obtained from in-air tests mainly at room temperature. Fatigue tests on large-scale carbon steel vessel performed in air at room temperature have shown that the cracks may initiate below the fatigue design curves, but that wall penetration is not expected until the fatigue cycles exceed the ASME design curves by about a factor of 3 (Cooper 1992). However, recent results from fatigue tests show that the effects of high-temperature pure water (simulating a PWR environment) on the fatigue strength (resistance to crack initiation) of carbon steels are not fully accounted for in the ASME design curves (Terrell 1988).

2.4 Vibratory Fatigue. Vibratory fatigue failures of piping have occurred predominantly at the socket welds commonly used for joining small diameter piping (< 50 mm) and at the fillet-welded attachments. In several instances, these failures have resulted in complete ruptures of piping. There are two approaches for high-cycle vibrational fatigue analysis of socket-welded small-diameter piping connections and fillet welded attachments: (1) ASME Section III fatigue design, which is an analysis-based approach, and (2) American Association of State Highway Transportation Officers fatigue design, which is an empirical approach.

2.5 Primary Water Stress Corrosion Cracking: Primary water stress corrosion cracking (PWSCC) has caused cracking, and in some cases, leakage from steam generator tubes and tube plugs, pressurizer instrument penetrations and heater sleeves, control rod drive mechanism nozzles, and reactor coolant piping penetrations. PWSCC is an intergranular cracking mechanism requiring at least the following three conditions to be present simultaneously: (1) high applied or residual tensile stress or both, (2) susceptible tubing microstructure (few intergranular carbides), and (3) high temperature. The PWSCC damage rate increases as a function of stress to an exponent, typically equal to 4. The PWSCC resistance of Alloy 600 is highest when the grain boundaries are covered with continuous or semicontinuous carbides. PWSCC is a thermally activated process and, therefore, its initiation and growth are very sensitive to temperature. These relationships are used in predicting the PWSCC initiation time in Alloy 600 components.

2.6 Flow-Accelerated Corrosion. FAC affects carbon steel piping, thermal sleeves, J-tubes, and feedrings carrying single phase, subcooled feedwater and steamlines carrying wet steam. FAC causes wall thinning and may lead to catastrophic failure of the affected components. Such failures have been reported in the field. This mechanism is further discussed in Section 3 and the available models are identified there.

2.7 Selection of a Degradation Mechanism for Incorporation into PRA. The degradation mechanisms may be divided in two groups based on the resulting failure modes: (1) those that may cause rupture, and (2) those that may cause cracking. Radiation embrittlement, thermal aging of cast stainless steel components, PWSCC and IGSCC of steam generator tubes, vibratory fatigue of small-diameter piping, and FAC may cause rupture, as mentioned earlier, whereas low-cycle fatigue, high-cycle thermal fatigue, and stress corrosion cracking of components other than steam generator tubes may cause cracking. The mechanisms that have potential to cause rupture are likely to have significantly more risk impact. Therefore, we have decided to incorporate a model for one of the degradation mechanism that may cause rupture.

We selected to incorporate FAC mechanism into PRA because, in some instances, it has caused rupture of feedwater or main steam line. In addition, depending on the affected piping, its risk impact can be significant. We did not select radiation embrittlement of reactor pressure vessels or stress corrosion cracking mechanisms acting on steam generator tubes because other USNRC projects have evaluated or are evaluating their risk-impact. We did not evaluate vibratory fatigue because the associated failures are caused by premature aging resulting from design and fabrication deficiencies.

3. Flow-Accelerated Corrosion

A brief description of FAC is as follows. A thin layer of porous iron oxide [mostly magnetite (Fe_3O_4)] forms on the inside surface of carbon steel feedwater piping exposed to deoxygenated water in the temperature range of about 95 to 260°C (200 to 500°F). Generally, this layer protects the underlying piping from the corrosive environment and limits further corrosion. However, the magnetite layer may be dissolved at the oxide-water interface and be replaced by new iron oxide formed at the metal-oxide interface, resulting in material removal and thinning of the piping. This process is called *single-phase FAC*. A similar corrosion process causes wall thinning of carbon steel piping exposed to wet steam; this process is called *two-phase FAC*.

The FAC phenomena follows a simple two-step process illustrated in Figure 1 (Remy and Bouchacourt 1992). The first step consists of production of soluble ferrous ions and their accumulation at the oxide-water interface and the second step consists of mass transfer of these ions into the bulk coolant. In the second step, the flowing water removes the soluble ferrous ions by a convective mass transfer mechanism, which is a diffusion gradient driven process. Generally, the concentration of ferrous ions in the bulk water is very low compared to their concentration at the oxide-water interface. Therefore, the ferrous ions present at the oxide-water interface can diffuse very rapidly into the solution.

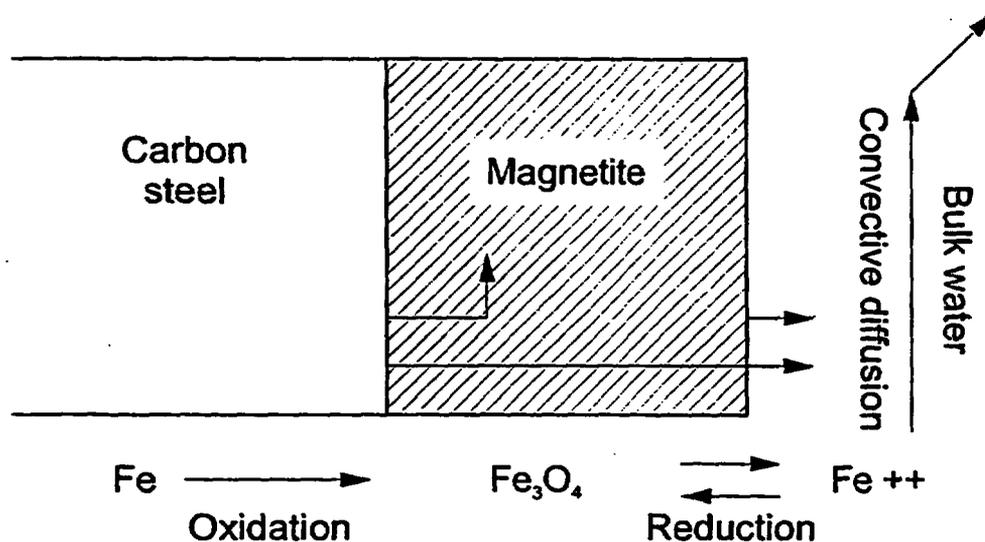


Figure 1. The flow-accelerated corrosion model (Remy and Bouchacourt 1992). C98 0936

3.1 Single-Phase Flow-Accelerated Corrosion. Single-phase FAC test results have identified several factors that affect the corrosion rate: (a) hydrodynamic variables - fluid velocity, piping configuration (geometry of the flow path), and roughness of the pipe inside surface; (b) metallurgical variables - chemical composition including weight percentage of chromium, molybdenum and copper in the steel; and (c) environmental variables - temperature and water chemistry including dissolved oxygen, ferrous ion concentration, pH, and amines used for pH control.

The *hydrodynamic variables* affect the rate of mass transfer of the iron ions and other corrosion products to the bulk coolant and thus affect the FAC rate. *Fluid velocity* affects the mass transfer. At a relatively low fluid velocity, the corrosion rate is controlled by the rate of mass transfer, whereas at higher velocity (still lower than the critical velocity above which metal removal by mechanical process takes place), the mass transfer rate is higher and the corrosion rate is controlled by the chemical reactions at the oxide-coolant and metal-oxide interfaces. FAC is less frequently observed in straight lengths of pipe free from hydrodynamic disturbances unless the bulk fluid velocity is high. Laboratory studies of the effect of bulk flow velocities, which varied from 2 to 18 m/s, on the corrosion of carbon steel in 150°C (300°F) circulating water show that the corrosion rate increases with an increase in the flow rate and, for a given flow rate, the corrosion rate is about constant.¹ The variable *piping configuration* takes into account the hydrodynamic disturbances (elbows, tees, reducers, valves, flow control orifices, etc.) that produce high local fluid velocities and result in a further increase in mass transfer. Experiments have shown that local-flow velocities in elbows can be two to three times bulk-flow velocities (Bosnak 1987, USNRC 1987a).

¹ Feedwater flow in a typical four-loop, 1100-MWe unit is about 6.8 x 10⁶ kg/h (15 x 10⁶ lb/h), which corresponds to about 4 m/s bulk flow velocity.

Chemical composition of carbon steel consisting of trace amounts of chromium, molybdenum, and copper provides resistance to FAC. The corrosion rate is most sensitive to the weight percent (wt%) of the chromium in steel. The corrosion rate is insignificant for Chromium content greater than 0.1 wt%.

Two main *environmental variables* that affect FAC rate are coolant temperature and water chemistry. The *coolant temperature* influences both the ferrous ion production and the mass transfer of these ions into the bulk water (Remy and Bouchacourt 1992). As the temperature increases, the ferrous ion concentration at the oxide-water interface decreases almost linearly. On the other hand, as the temperature increases, the ferrous ion diffusivity into the coolant increases, resulting in a mass transfer coefficient that increases about linearly. The resulting corrosion rate variation with temperature is a bell-shaped curve. For the typical conditions in the PWR feedwater piping, the maximum corrosion rate occurs at about 150°C (300°F) (Chexal and Horowitz 1995). The *water chemistry* includes dissolved oxygen, ferrous ion concentration, metallic impurities, and cold pH level. The FAC rate varies inversely with the level of *dissolved oxygen* in the fluid. As the level of oxygen increases above a threshold value, a less porous oxide layer of hematite, instead of magnetite, is formed. Because the solubility of hematite in the feedwater is several orders of magnitude lower than that of magnetite, the FAC rate decreases significantly. Laboratory test results show that the threshold value for dissolved oxygen is less than 15 ppb (Remy and Bouchacourt 1992).

Ferrous ion concentration and metallic impurities in the water affects the FAC rate. The increase in the *ferrous ion concentration* in the bulk fluid reduces the mass transfer of ferrous ions from the oxide-coolant interface to the bulk coolant. An increased ion concentration can reduce or suppress FAC when the process is controlled by mass transfer. FAC rates vary by an order of magnitude over the cold *pH range* of 8.5 to 9.5, which is typical for PWR feedwater systems (Shack and Jonas 1988).

3.2 Two-Phase Flow-Accelerated Corrosion. Examination of worn extraction piping has identified two distinct *mechanisms* causing damage in the system carrying two-phase coolant: oxide dissolution and droplet-impact wear (Bosnak 1987, Keck and Griffith 1987). The oxide dissolution mechanism is similar to single-phase FAC mechanism discussed with one exception. Two-phase FAC has been observed in piping carrying wet steam. Its occurrence has not been observed in piping carrying dry steam (100% quality). Moisture in the wet steam is essential to dissolve the oxide film. Field data indicate that the greatest degradation is seen in the piping containing steam with the highest moisture content, such as the turbine crossover piping and the exhaust and extraction piping connected to the high-pressure turbines.

The droplet-impact wear mechanism, as its name implies, cause damage by mechanical action. The impact of liquid droplets, entrained in steam, on carbon-steel oxide films can produce a matrix of cracks and subsequent fatigue failure of the films, and expose the underlying metal surfaces to the corrosive action of the coolant. (Keck and Griffith 1987).

3.3 Computer Models. EPRI has developed the computer code CHECWORKS (Chexal-Horowitz Engineering-Corrosion Workstation) for managing FAC of nuclear power plant piping. This program has capabilities for estimating parameters (such as local water chemistry and flow rate) that affect corrosion rates, and for predicting relative corrosion rates to facilitate selection of inspection locations. The computer code is based on laboratory data from France, England, and Germany, and on U.S. plant data. The code has been validated using other U.S. plant data. The main sources of uncertainties are associated with the original thickness and thickness profile of the piping components, trace amounts of alloy content in the piping

material, actual number of hours of operation, plant chemistry history, and geometric discontinuities on the inside surface of the piping.

All PWR and BWR plants in the U.S. use the CHECWORKS code (or its predecessor code CHECMATE) for estimating the FAC rates (Chexal and Horowitz 1995). This code is also used by many fossil plants, by the U.S. Navy, and by several overseas utilities. The code has been used for identifying the sites most susceptible to FAC, prioritizing the locations for inspection, estimating remaining service life for susceptible components, and evaluating the effectiveness of different water chemistries and other mitigative actions.

Siemens/KWU in Germany has long been active in researching wall thinning rate estimation caused by FAC. The empirical Kastner model was developed in the early 1980s for the calculation of material losses due to FAC in single- and two-phase flow (Kastner and Riedle 1986). This model is hereafter referred to as the KWU-KR model. The KWU-KR model is based on experiments carried out at Siemens/KWU and on plant data from all the known single and two-phase locations (more than 6,000 data points overall), as well as on theoretical considerations. After the Surry Unit 2 accident in 1986, the WATHEC program based on the KWU-KR model was developed to perform weak-point analyses at power plants. In 1991, WATHEC was interfaced with the DASY program which handles the recording, management, evaluation, and documentation of the data obtained from non-destructive examinations. These two software packages were continuously improved in cooperation with European utilities to calibrate the predicted wall thinning rates for further plant diagnosis with increased prediction accuracy (Chexal et al. 1996).

The FAC models discussed here are developed and validated to estimate the relative corrosion rate at different locations in the piping and not to predict pipe rupture, which is needed for assessing safety impact. Simplified assumptions have been made to estimate the capacity of degraded piping. These assumptions are discussed in Section 5.

We have elected to base our calculations on the KWU-KR model because it is well documented in the published literature. Similar documentation for the CHECWORKS model is not available because of its proprietary nature. A third model, part of the BRT-Cicero code, is based on the test data taken at the Cicero test loop and was developed by the Electricité de France. But the documentation of this model is also not available because of its proprietary nature.

3.4 The KWU-KR Model. This model calculates the corrosion rate as a function of Keller's geometry factor, flow velocity, fluid temperature, material chemical composition, fluid chemistry (pH at 25°C and dissolved oxygen), exposure time, and, in the case of two-phase flow, steam quality. With this model, the pipe thickness can be calculated as a function of time. The wall corrosion, $W_c(t)$, is the thickness of the pipe corroded away and is calculated by

$$W_{C,calculated}(t) = \frac{\Delta\phi_R t}{\rho_{st}} \quad (1)$$

where

| | | |
|----------------|---|---|
| $\Delta\phi_R$ | = | FAC rate ($\mu\text{g}/\text{cm}^2 \text{ hr}$), |
| t | = | exposure time (hr), |
| ρ_{st} | = | the density of steel ($\mu\text{g}/\text{cm}^3$). |

Kastner and Riedle's work (1987) is the basis for estimating the FAC rate, $\Delta\phi_R$. Once the FAC rate is estimated, the wall thickness as a function of time can be calculated. This wall thickness is found by

$$W_{pipe}(t) = W_{original} - W_{C,calculated}(t) \quad (2)$$

where

$$\begin{aligned} W_{pipe}(t) &= \text{pipe wall thickness at time } t \text{ (cm),} \\ W_{original} &= \text{original, nominal pipe wall thickness (cm),} \\ W_{C,calculated}(t) &= \text{thickness of pipe corroded away at time } t \text{ (cm).} \end{aligned}$$

The corrosion rate, $\Delta\phi_R$, is calculated via the following steps.

1. Using the KWU-KR model, the pH, oxygen content, liquid velocity, geometrical factor, total content of chromium and molybdenum in steel, and operating temperature are known. We can calculate FAC rate as the following equations [Kastner, 1986]:

$$\Delta\phi_R = 6.35 k_c (B \cdot e^{N \cdot w} \cdot [1 - 0.175 \cdot (pH - 7)^2] \cdot 1.8 \cdot e^{-0.118 R} + 1) \cdot f(t) \quad (3)$$

with

$$\begin{aligned} B &= -10.5\sqrt{h} - (9.375 \times 10^{-4} T^2) + (0.79 T) - 132.5 \\ N &= -0.0875 h - (1.275 \times 10^{-5} T^2) + (1.078 \times 10^{-2} T) - 2.15 \text{ (for } 0\% \leq h \leq 0.5\%) \\ N &= (-1.29 \times 10^{-4} T^2 + 0.109 T - 22.07) 0.154 e^{-1.2h} \text{ (for } 0.5\% \leq h \leq 5\%) \end{aligned}$$

where

$$\begin{aligned} \Delta\phi_R &= \text{calculated specific rate of material loss } (\mu\text{g/cm}^2 \text{ h}), \\ k_c &= \text{geometrical factor,} \\ w &= \text{flow velocity (m/s),} \\ pH &= \text{pH value,} \\ g &= \text{oxygen content } (\mu\text{g/kg}), \\ h &= \text{content of chromium and molybdenum in steel (total \%),} \\ T &= \text{temperature } (^\circ\text{K}). \\ f(t) &= \text{a time correction factor} \end{aligned}$$

Note that the time correction factor, $f(t)$, of the FAC rate equation is a function of exposure time in the KWU-KR model [Kastner and Riedle, 1986]. Exploring the behavior of this factor, it can be shown that the factor $f(t)$ has a value of 1 in small operating periods and tends to the value of 0.79 for an operating period of 9.6×10^4 hrs (around 11 years). For longer periods ($t \geq 9.6 \times 10^4$ hrs), $f(t)$ equals 0.79. The time correction factor is given by

$$f(t) = C_1 + C_2 t + C_3 t^2 + C_4 t^3 \quad (4)$$

where

| | | |
|----------------|---|---------------------------------|
| t | = | the exposure time (hr), |
| C ₁ | = | 9.999934 × 10 ⁻¹ , |
| C ₂ | = | -3.356901 × 10 ⁻⁷ , |
| C ₃ | = | -5.624812 × 10 ⁻¹¹ , |
| C ₄ | = | 3.849972 × 10 ⁻¹⁶ . |

The geometry factor, k_g, is given by one of the following values:[Kastner, 1986]

| | |
|---|---|
| 0.04 for "straight tube" | 0.08 for "leaky joints" "labyrinths" |
| 0.15 for "behind junctions" | 0.16 for "behind tube inlet (sharp edge)" |
| 0.23 for "elbow R/D=2.5" | 0.30 for "elbow R/D=1.5" |
| 0.30 for "in and over blades" | 0.52 for "elbow R/D=.5" |
| 0.60 for "in branches #2" | 0.75 for "in branches #1" |
| 1.0 for "on tubes" "on blade" or "on plate" | |

This FAC model was developed in 1980s and, therefore, the further understanding developed since then is not incorporated in the model. The following assumptions are employed in the model. Comments based on the current understanding of the FAC phenomena are also presented, if appropriate, along with each assumption.

1. The model has no restriction on the flow velocity up to the critical velocity which metal removal takes place by mechanical processes.
2. The FAC rates are insignificant at water temperature greater than 240°C, and the resulting material losses can be ignored. This assumption is consistent with our current understanding.
3. The lower and upper limits for the cold pH are 7.0 and 9.39, respectively. Note that the typical limits for PWR are 8.5 and 9.5. The model assumes that the corrosion rate is very small (1 μg/cm²/h) and constant if the pH is greater than 9.39. But the test results show that at higher pH values, the corrosion rate increases with increasing pH value.
4. The oxygen concentration is less than 30 ppb. For higher concentration, the rate is constant and very small. The test results and plant data show that the rate is very small for the oxygen concentration greater than 15 ppb.
5. The chromium and molybdenum content is less than 0.5 wt%. No material loss takes place if the content is higher. This is conservative because the field data show that there is no material loss if the content is greater than 0.1 wt%.
6. The model is valid only for operating periods longer than 200 h. Very high losses can occur in the start-up phase.
7. The two-phase FAC model uses the mean velocity in the water film on the inside surface of the piping instead of the velocity of a two-phase fluid.

4. How Flow-Accelerated Corrosion Can Impact Risk

FAC has caused rupture of carbon steel secondary piping at several PWRs. The most notable rupture of feed water piping occurred at Surry, Unit 2, on December 9, 1986. Another noteworthy event associated with single-phase FAC is the rupture of a drain pump discharge piping (350-mm diameter) at Trojan on March 9, 1985. A pressure transient caused the ultimate rupture of feedwater piping already significantly degraded by FAC at both plants. In neither case were there a leak or any other warning signs indicating incipient failures.

Two examples of events associated with two-phase FAC include ruptures of the fourth stage steam extraction piping, one at Millstone Unit 2 in October, 1986, and another one at Fort Calhoun on April 21, 1997. Single-phase FAC has also caused significant wall thinning in auxiliary feedwater piping at Catawba Unit 2, which has preheat steam generators. Ten Westinghouse-designed PWRs with Models D4, D5, and E of preheat steam generators are susceptible to such wall thinning (USNRC 1992a). The pipe rupture is a risk-significant event because it may cause reactor trip. In addition, some of the rupture events, such as a break of a small steam line, may represent a dominating overcooling event that may contribute to pressurized thermal shock (PTS) risk to reactor pressure vessel (Selby et al. 1985). The PTS risk associated with steam line break is, however, small; therefore, the PTS risk associated with FAC-induced steam line rupture will be quite small.

All the FAC-related field experience in PWR plants is associated with piping outside the containment. But single-phase FAC has also caused significant wall thinning of ceratin carbon steel piping located inside the containment at one BWR (USNRC 1992b). Since this piping constitutes a primary pressure boundary, the event has raised a concern that the rupture caused by FAC may lead to a loss-of-coolant accident. In general though, FAC may have following four types of risk impacts:

1. May increase the frequency of transients.
2. May increase the unavailability of safety system.
3. May contribute to LOCA in certain BWR plants.
4. May increase the frequencies of certain overcooling events responsible for generating pressurized thermal shock.

5. Incorporation of Flow-Accelerated Corrosion Model into PRA

To determine the risk impact of pipe rupture caused by FAC, we need to determine the probability of rupture at different locations in susceptible piping. The probability of rupture can be determined using the load-capacity formulation for the FAC mechanism. The loads acting on the piping may be represented by the steady state and transient pressures, and the pressure capacity of piping takes into account pipe wall thinning caused by FAC. We first describe general load-capacity formulation and then present details for developing probability density functions for load and capacity.

5.1 General Load-Capacity Formulation. Let the pressure capacity be C (in our case, the system maximum allowable pressure is a function of time due to the thickness of the pipe that has been corroded away). Let the stress or load be L (i.e., the static load from the system operation pressure at full power or steady state operation and the dynamic load during a pressure transient). The probability of failure of the piping can be computed from

$$P_f(t) = P[C(t) < L(t)] = 1 - \int_0^{\infty} F_L(x, t) f_C(x, t) dx$$

$$= \int_0^{\infty} \left(\int_x^{\infty} f_L(y, t) dy \right) f_C(x, t) dx \quad (5)$$

where

| | | |
|-------------|---|--|
| $P_f(t)$ | = | pipe failure probability, |
| $C(t)$ | = | pipe pressure capacity (ksi), |
| $L(t)$ | = | pipe pressure load (ksi), |
| $F_L(x, t)$ | = | the load cumulative distribution function, |
| $f_C(x, t)$ | = | the capacity probability density function (1/ksi), |
| $f_L(y, t)$ | = | the load probability density function (1/ksi), |
| t | = | operational time (hr), |
| x, y | = | capacity and load pressures, respectively (ksi). |

Equation 5 is based on the simple observation that failure occurs when the load exceeds the capacity. The capacity probability density function, $f_C(x, t)$ can be determined from the remaining pipe wall thickness, $W_{pipe}(t)$, as defined in Equ. (2). The KWU-KR model for the corroded wall thickness, $W_{C,calculated}(t)$, has already been defined in Equ. (1). The pipe pressure capacity evaluation would then incorporate the thinning wall (which is a function of time) to determine the expected pressure capacity. To perform this evaluation, we made the assumption that the wall thinning is uniform around the circumference. Then, the failure pressure as a function of time can be calculated from the following equation given by Wesley et al. (1990).

$$ps_f = \frac{\sigma_f \cdot W_{pipe}(t)}{[r + W_{C,calculated}(t)](1 + 0.25 \epsilon_f)} \quad (6)$$

where

| | | |
|-----------------------|---|---|
| $ps_f(t)$ | = | failure pressure (ksi), |
| σ_f | = | failure stress (ksi), |
| $W_{pipe}(t)$ | = | pipe wall thickness at time t (cm), |
| r | = | initial inside radius (cm) |
| $W_{C,calculated}(t)$ | = | thickness of pipe corroded away at time t (cm) [Equ. (1)] |
| ϵ_f | = | median hoop strain at failure (failure strain). |

The above equation defines the failure pressure for a straight pipe in terms of hoop stress and includes some provision for the biaxial stress-state in pipe wall and strain concentration effects. Failure stress and hoop strain at failure are determined from 51-mm (2-in.) uniaxial tensile test specimens. Both the failure stress and strain should be treated as random quantities because of variability in the stress-strain relationship,

uncertainty related to biaxial stress condition, necking, and effective gage length. Mean values for failure stress, σ_f , and failure strain, ϵ_f , for SA 516 grade 70 carbon steel are shown below in Table 1 (Wesley et al. 1990). Although the above equation is for a straight pipe, we have assumed that it is applicable to pipe fittings such as elbows.

Table 1. Typical failure stress and strain (hoop strain) values for SA 516 grade 70 carbon steel.

| Temperature (°F) | Failure stress (σ_f), ksi | Hoop strain (ϵ_f), % |
|------------------|------------------------------------|---------------------------------|
| 77 | 75.6 | 6.2 |
| 400 | 78.3 | 3.7 |
| 600 | 76.5 | 5.8 |
| 800 | 63.9 | 7.9 |

The load distribution, $f_L(y, t)$, represents the actual and anticipated pressures for a particular section of piping. Normally, the majority of this distribution will be at the nominal system operational pressure. For example, if nominal system pressure operation is approximately 900 psig and is experienced 95% of the time during operation, the load distribution would have 95% of the distribution centered on or around 900 psig. Transient pressures that cause system pressure to exceed nominal pressures would need to be incorporated into the load distribution. But, transient pressures are represented by an aleatory type of model. Incorporating these pressures would entail determining the anticipated pressures and likelihood of experiencing such a pressure as a function of time.

Once both the load and capacity distributions are known, we can estimate the piping failure probability. To perform this estimation, the density functions for the distributions could be utilized to obtain an analytic expression for the pipe reliability as a function of time. This approach was used for the results in this paper. An alternative approach is to utilize a Monte Carlo simulation routine to determine the fraction of time that the load pressure is larger than the capacity pressure. This fraction would then directly represent the failure probability of the pipe due to FAC.

Looking at the "top-most" modeling level for the pipe segment, the load-capacity FAC model² is comprised of two parts. First, the *deterministic* aspect is included and is numerically determined by using the pipe wall thickness [$W_{pipe}(t)$] calculation as described in Section 3. A second, and equally important part of the load-capacity model, is an aleatory model representing the arrival of transient overpressure loadings (where "overpressure" indicates a pressure transient above the nominal steady-state pressure). Beneath these top-level models, we introduce applicable *epistemic* uncertainties (Apostolakis 1995). As illustrated in Figure 2, under the deterministic portion of the model are epistemic uncertainties of two types, model and parametric. Specifically, these include the E factor (model); the fluid parameters, including the nominal, steady-state pressure (parametric); and the piping parameters. Under the aleatory portion of the model are epistemic uncertainties of a single type, parametric. The parametric uncertainty in this case refers to the rate of occurrences of transient pressures. These aspects of the load-capacity model are clearly delineated since they will dictate how the model is solved as part of the PRA.

² The "model of the world" in the terminology of Apostolakis (1995).

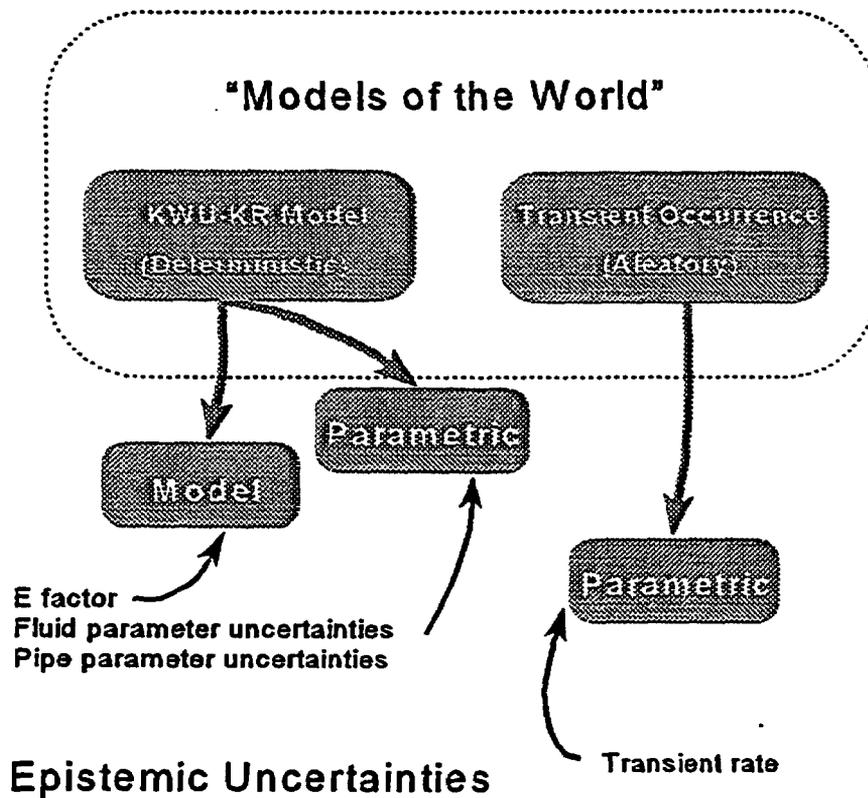


Figure 2. Graphical illustration of the load-capacity FAC model attributes and uncertainty types.

5.2 Detailed Determination of Capacity. The capacity probability density function, $f_c(x, t)$, can be determined using Equ. (4), (1), and (2) along with the KWU-KR model for $\Delta\phi_R$. The uncertainties in the quantities on the right-hand side of Equ. (4) determine $f_c(x, t)$. Of these, the most important is the uncertainty in the prediction of the KWU-KR model itself.

From the published literature [Kastner and Riedle, 1986], the relationship between the empirical model predictions and the measured FAC rates in laboratory studies or in power plants can be determined. This relationship is shown in Figure 3. From the data, it appears that the KWU-KR model has been developed to over-predict the FAC rate. In other words, it is designed to err mostly on the conservative side and is typical of models utilized for nuclear power plant safety where conservatism is *de rigueur*. But, this feature of the model can be used to determine an "adjustment" factor that would express our uncertainty in the calculated results. Within Figure 3, there are a total of 1,049 cases where the variability reflected on individual data points is due to both parameter uncertainties and the uncertainty due to the model itself.

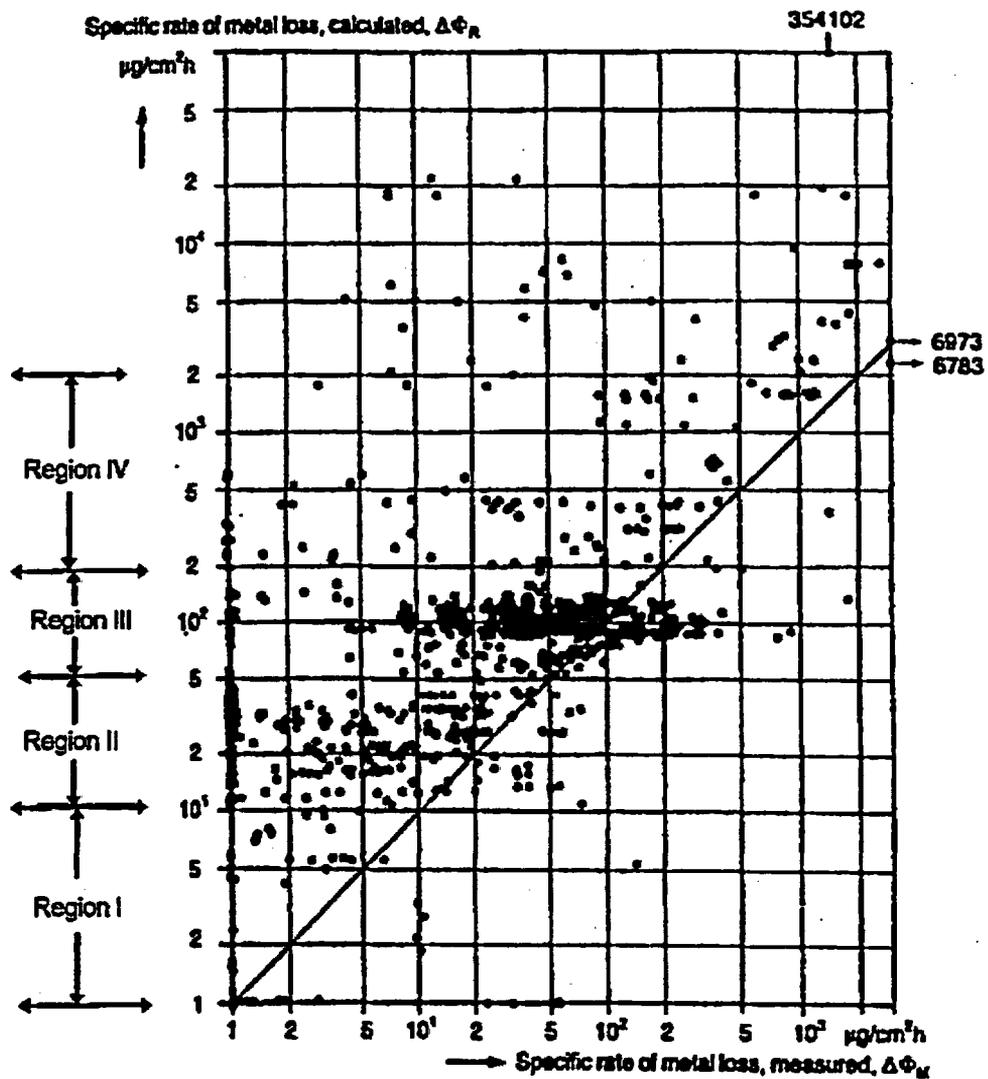


Figure 3. Comparison of values calculated by empirical KWU-KR model with measurements from laboratory experiments and power stations. [adapted from Kastner and Riedle, 1986]

NOTE:

Region I: 1 to 10 $\mu\text{g}/\text{cm}^2\text{h}$
 Region III: 50 to 200 $\mu\text{g}/\text{cm}^2\text{h}$

Region II: 10 to 50 $\mu\text{g}/\text{cm}^2\text{h}$
 Region IV: 200 to 2,000 $\mu\text{g}/\text{cm}^2\text{h}$

To express our uncertainty in the KWU-KR model predictions, we employ the "adjustment-factor" approach discussed by Siu and Apostolakis (1982) and Apostolakis (1995). This approach is as follows:

We take the available model, the so called deterministic reference model, and introduce a multiplicative factor E to modify its results. Referring to Equ. (1), the FAC rate is expressed as

$$\Delta\phi_R = \Delta\phi_{R, KWU-K} \times E \quad (7)$$

where

- $\Delta\phi_R$ = the specific FAC rate to be used in Equ. (1) ($\mu\text{g}/\text{cm}^2 \text{ h}$),
- $\Delta\phi_{R, KWU-K}$ = the specific FAC rate as predicted by the KWU-KR model (the deterministic reference model) ($\mu\text{g}/\text{cm}^2 \text{ h}$),
- E = the adjustment factor.

Thus, Equ. (1) becomes:

$$W_{C, calculated}(t) = \frac{E (\Delta\phi_{R, KWU-K}) t}{P_{st}} \quad (8)$$

We note that Equ. (5) indicates that the actual specific FAC rate can be considered to be the product of its deterministic reference model prediction and an adjustment factor E which accounts for the inadequacy of the calculated value. The question is now what type of uncertain variable is E. To determine this, we look at the evidence contained in Fig. 3.

If there were no model uncertainties in the KWU-KR model, the points in Fig. 3 would fall on the straight line with a slope of one. This would indicate that the calculated (on the ordinate) and the actual (on the abscissa) values of the specific FAC rate were the same. Clearly, this is not the case. There are a number of measured values that correspond to a given calculated value. Presumably, the authors of the paper in which this figure appeared originally made every effort to ensure that the parameter values in the experiments and the power stations from which the measured values were taken were very close to the parameter values that were input to the KWU-KR model to generate the calculated values. In other words, the discrepancies should be primarily due to model uncertainty.

When we make predictions using the KWU-KR model, we do not know how inaccurate the model is with respect to the actual circumstances that we are attempting to model, even though the parameter values (pH, flow rate, etc.) that are input to the model are (almost) the same as those anticipated in the actual circumstances. In other words, we do not know what the value of E (the ratio of the actual over the calculated value of the specific FAC rate) is. We can say, then, that we have a population of circumstances similar to Kaplan's (1983) population of plant-specific failure rates. This "circumstance variability" is aleatory, i.e., the particular circumstances of interest to us will be one of the many circumstances shown in Fig. 3 (always for the same parameter values). The relative frequency of the points with the same calculated value of the specific FAC rate determines the likelihood that the factor E will have the corresponding value. This variability will be there as long as we use the KWU-KR model (as described by Kastner and Riedle, 1986) to produce the calculated value and it is due to the approximations made to develop this model, i.e., it is an aleatory variability.

The recognition that there is "circumstance variability" leads us naturally to the idea of a "two-stage" Bayesian analysis, as Kaplan (1983) has recommended. Siu and Apostolakis (1985) have, in fact, proposed an approach to estimate the distribution of E when evidence becomes available. The aleatory probability distribution of the factor E is assumed to be lognormal with parameters μ and σ . The epistemic uncertainty is, in this case, described by a probability density function over the parameter vector $[\mu, \sigma]$. Each value of this vector specifies one aleatory distribution for E. The average of these aleatory curves is used in the second stage as the epistemic distribution of E for a specific set of circumstances.

In the present case, we note that Fig. 3 contains a fairly large number of points. Thus, the "two-stage" calculations (and the attendant calculational complexity) are not really needed, since the average aleatory distribution of E can be derived by simply using standard software packages to fit distributions to data. But, we cannot do this for every possible value of the calculated FAC rate. Instead, we divide the range of the calculated values of specific rates of metal loss in Fig. 3 into four regions (denoted Region I through Region IV). Points above the calculated specific rates of metal loss of $2,000 \mu\text{g}/\text{cm}^2\text{h}$ and points below the measured rate of $1 \mu\text{g}/\text{cm}^2\text{h}$ are excluded from the calculations.

The lognormal distribution is found to fit the data in each region very well. The parameters μ , σ of the distribution, as well as several characteristic values, are shown in Table 2.

Table 2. Lognormal distributions for the E factor for the four regions of Fig. 3.

| | Region I | Region II | Region III | Region IV |
|-----------------------------|----------|-----------|------------|-----------|
| 5 th percentile | 0.20 | 0.05 | 0.09 | 0.009 |
| 50 th percentile | 1.39 | 0.31 | 0.49 | 0.13 |
| 95 th percentile | 9.47 | 2.06 | 2.66 | 1.99 |
| Mean | 1.43 | 0.61 | 0.83 | 0.51 |
| Error Factor | 6.83 | 6.62 | 5.46 | 14.99 |
| μ | 0.33 | -1.17 | -0.72 | -2.03 |
| σ | 1.17 | 1.15 | 1.04 | 1.65 |

As an example, suppose that we perform our calculations with the KWU-KR model and the result is $25 \mu\text{g}/\text{cm}^2\text{h}$. The actual value of the specific FAC rate will be [see eq. (5)]

$$\Delta\phi_R = \Delta\phi_{R, KWU-KR} * E_{II} \quad (9)$$

where E_{II} is the factor corresponding to Region II. The actual specific FAC rate will also be lognormally distributed with parameters

$$\mu_{\Delta\phi_R} = -1.169 + \ln(\Delta\phi_{R, KWU-KR}) = 2.050 \quad (10)$$

and

$$\sigma = 1.154 \quad (11)$$

Several characteristic values of the actual specific FAC rate are ($\mu\text{g}/\text{cm}^2\text{h}$):

| | | | |
|-----------------------------|------|-----------------------------|------|
| 5 th percentile | 4.1 | 50 th percentile | 7.8 |
| 95 th percentile | 14.8 | Mean | 15.1 |

We observe that the evidence shown in Fig. 3 leads to the conclusion that, even though the calculated value of the specific FAC rate is $25 \mu\text{g}/\text{cm}^2\text{h}$, the 95th percentile of the *actual* rate is only $14.8 \mu\text{g}/\text{cm}^2\text{h}$. It is evident from Fig. 3 that the KWU-KR model predictions are conservative.

5.3 Detailed Determination of the Load. As discussed in Section 4, the "load" in the load-capacity FAC model represents the actual operating pressures that will be seen by the pipe undergoing the FAC process. In our nomenclature, the load distribution, $f_L(y, t)$, is given by two different types of pressures.

1. Nominal, steady-state pressures
2. Transient pressures

These two loads are treated separately and with different model types. The steady-state pressure is modeled simply as a parameter with epistemic uncertainty. But, the treatment of the transients is more complicated because their occurrence in time must be included in Equ. (3). This occurrence of transients is, of course, an aleatory phenomenon that is modeled using the Poisson distribution. A series of calculations leads to the following expression for the reliability of a component

$$R(t_L) = \exp\left(-\lambda t_L \left[1 - \frac{1}{t_L} \int_0^{t_L} F_L[C(t_0) g(t)] dt\right]\right) \quad (12)$$

We note that this is the aleatory model for the reliability, i.e., this is the model of the world for transients (Apostolakis 1995). The predictive reliability will be the average over the epistemic distribution of the initial capacity, i.e.,

$$\overline{R(t_L)} = \int_0^{\infty} \exp\left(-\lambda t_L \left[1 - \frac{1}{t_L} \int_0^{t_L} F_L[C(t_0) g(t)] dt\right]\right) f_{C_0}(x) dx \quad (13)$$

where

$$f_{C_0}(x) = \text{the probability density function of the initial capacity.}$$

Note that the probability density function of the initial capacity is determined primarily by the epistemic uncertainty in the factor E. Further, note that an alternative procedure to the calculation given by Equ. (12) would be to simulate the occurrence of transients (and then calculate the FAC rate for that particular time period). The treatment of the transient pressures is more difficult than the steady-state pressures since the *time* to an transient overpressure is a random variable. But, either of the two methods will provide an appropriate analysis method.

Finally, the probability of failure due to transients is the complement of Equ. (12). The results of this calculation is shown in the next section.

6. Case Study

For our case study, a piping segment was selected for analysis and incorporation into a full-scale PRA model. The selected pipe segment in our analysis is the same as the particular segment that failed in December of 1986 (13.6 years after commercial operation) at Surry. The 18-inch suction line to the main feedwater pump A of Unit 2 failed in a catastrophic manner. The condensate feedwater system flows from a 24-inch header to two 18-inch suction lines, each of which supplies one of two feedwater pumps. In accordance with our choice to analyze the main feedwater (MFW) system, we specifically model the single pipe segment (FW-04) failure due to FAC in Surry's MFW system because this event also caused two main feedwater pumps to lose suction heads at the same time. Thus, we may assume that the event was a loss of the MFWs (corresponding to the IE-T2 initiating event in the Surry-IPE) which caused a plant transient (Virginia, 1991).

The indirect impacts due to failure of the "FW-04" pipe segment were also considered. For the Surry plant, flooding in the turbine building (91% contribution to core damage frequency) was a significant plant vulnerability. Flooding may occur as a result of failures in the circulating water and service water systems in the turbine building. Both of these water systems are gravity fed from the intake canal (20 feet above the Turbine Building basement floor) [Virginia, 1991]. The "FW-04" pipe segment is located in the basement of the turbine building. After discussions with Surry PRA and piping personnel, it was determined that there are no safety related equipment in the area. Hence the indirect effects do not contribute to our case study.

Once the FW-04 piping event was integrated into the Surry PRA model, the basic event was linked to the FAC "compound" plug-in calculation. With the data for the FAC parameters known, the failure probability of FW-04 as a function of time can be calculated. Figure 4 illustrates the results of such a calculation. Note that in addition to the total FW-04 failure probability, the constituents to the total probability are shown. As previously discussed, these constituents come from two type of models, namely, from either the nominal steady-state pressure or transient pressures.

The core damage probability due to FAC for 10 years shows a small increase for the loss of main feedwater event tree sequences (i.e., just the IE-T2 sequences). The total increase for these sequences over the nominal was found to be about 30%. However, the impact on the overall core damage frequency over the 10 year period was found to be insignificant (< 1%). This result is not surprising since the risk increase ratio (or risk achievement worth) importance measure for loss of MFW at the Surry plant has a value of only 1.2. Looking at the nominal loss of MFW sequences, we see that their contribution to the total core damage frequency is less than 1%. Note though that the impact on the risk due to FAC for other plants or plant models has not been evaluated and needs further investigation.

Our results can be used to check the validity of the linear failure rate model (Vesely, 1987). This model asserts that the failure rate of a component is a linear function of time, i.e.,

$$\lambda(t) = \lambda_0 + \alpha t \quad (14)$$

The term " αt " accounts for (potential) aging effects.

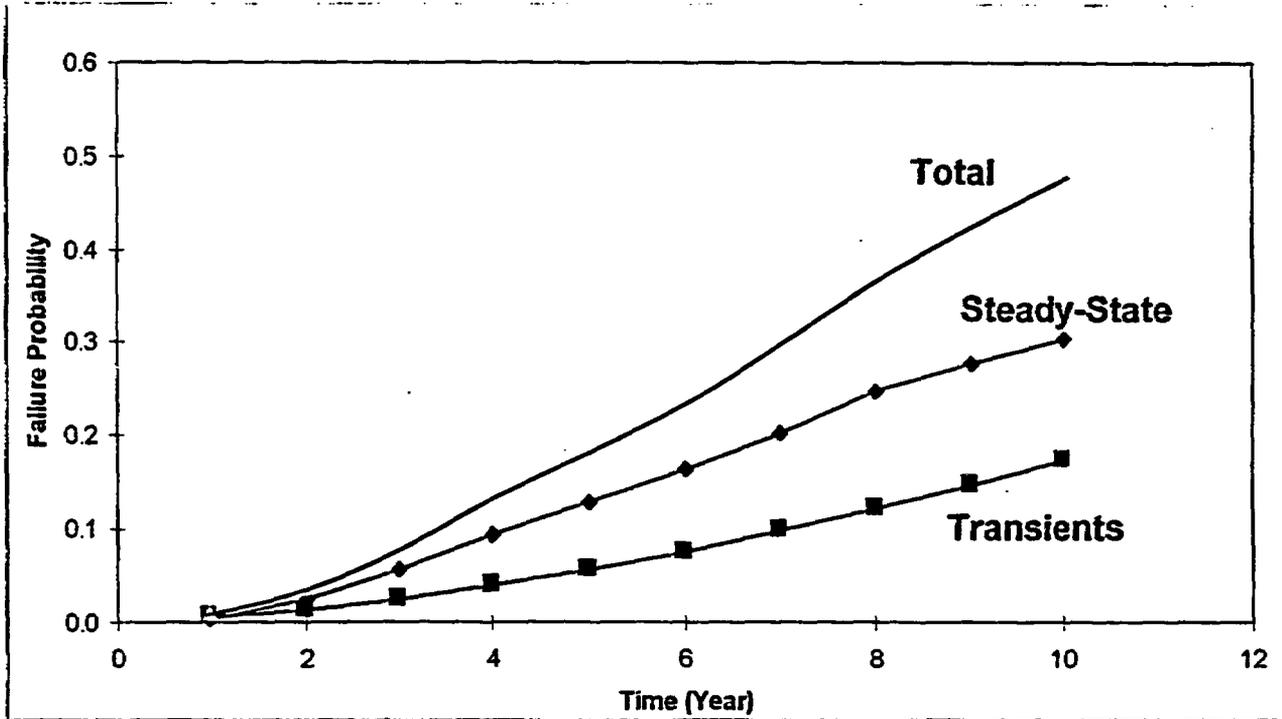


Figure 4. Probability of FW-04 pipe failure (due to FAC) as a function of time.

We can calculate the failure rate due to transients using the standard definition of failure rate, i.e.,

$$\lambda(t) = -\frac{dR(t)}{R(t)dt} \quad (15)$$

where $R(t)$ is the reliability of Equ. (12), i.e., the probability of surviving the transients.

We note that this failure rate, Equ. (15), must be calculated within the model of the world. In other words, it must be calculated *before* the averaging process of Equ. (13) is performed (equivalently, the failure rate can not be calculated directly from the curve corresponding to transients shown in Fig. 4). The average failure rate can be found from the predictive calculation

$$\overline{\lambda(t)} = \int_0^{\infty} \lambda(t) f_{C_0}(x) dx \quad (16)$$

Our results show that the assumption that the failure rate (due to transients) is linear with time is reasonable, i.e., Equ. (15) leads to linear $\lambda(t)$. The (epistemic) average value of α (from the case study) is equal to about $3 \times 10^{-3}/\text{yr}^2$. Consequently, the linear model would be a reasonable model to use in this case (assuming, of course, that the value of α could be obtained by other means, e.g., by expert opinion elicitation).

The model of the world for static pressures is deterministic, i.e., failure occurs when the capacity has deteriorated below the static pressure. Consequently, a failure rate similar to Equ. (15) does not exist. We note that it would be wrong to calculate a failure rate using the curve labeled "steady state" in Fig. 4, just as it would be inappropriate to calculate the average failure rate for transients from the curve corresponding to transients, as discussed above.

The conclusion is that, if one were to use the linear model of Equ. (15), assuming that α could be determined by methods other than those of reliability physics, one would model the probability of failure due to transients but would leave out the probability of failure due to static loads (because these are not aleatory and the concept of a failure rate does not exist in this case).

7. Conclusions

The overall objective of this work was to assess the *feasibility* of applying the "LANL/ASCA" method for incorporation of reliability-physics based models, expert judgement, and the results of the Nuclear Plant Aging Research (NPAR) program into an integrated aging risk assessment. The "LANL/ASCA" method refers to the application of PRA methods described in NUREG/CR-6157, *Survey and Evaluation of Aging Risk Assessment Methods and Applications* (Sanzo et. al., 1994). We successfully applied an aging reliability-physics model to the existing SAPHIRE Surry IPE risk model and subsequently determined aging-based risk insights.

As part of the work, general methods were developed to facilitate the inclusion of aging mechanisms into PRA models. The outcome of development included (1) a generic process to perform the aging-to-PRA modeling, (2) a focused application of FAC failure modeling, and (3) incorporation of the FAC aging model directly into a PRA. In summary, the general methods addressed the areas of:

- The technique of screening for risk-based determination of potential aging-affected components.
- The treatment of uncertainties (aleatory and epistemic).
- The integration of deterministic models into the PRA that represents applicable aging failure mechanisms.
- The determination of component aging-caused failure probabilities via a "load-and-capacity" analysis.

Again, since this is a feasibility study, it was not the objective of the project to address all aging mechanisms nor to provide a complete picture into the magnitude of the risk or core damage frequency impacts resulting from aging. The work did qualitatively discuss the spectrum of aging-related issues facing LWRs. Also, the work did evaluate aging impacts on the core damage frequency. But to reiterate, these evaluations are of a feasibility nature and *should not* be construed to represent the magnitude (both absolute and relative) of risk posed by aging in LWRs.

An important aspect of the work was the determination and management of key parts of the aging mechanism failure model. For the FAC failure model, a combination of deterministic and aleatory modeling was required. The deterministic portion of the failure model was provided by using the KWU-KR FAC rate model. The KWU-KR model was used to specify the piping capacity as a function of time. The loading on the piping was provided by utilizing two models, one deterministic (for steady-state pressures) and the other

aleatory (for transient pressures). Epistemic uncertainties on the model parameters, as well as the KWU-KR model itself, were treated and incorporated directly in the analysis. It is envisioned that, in general, other aging-related reliability-physics models must be constructed in a similar fashion.

We demonstrated that a rigorous treatment of an important aging mechanism such as FAC need not rely on sparse failure data in an attempt to quantify a statistical failure rate. Even though the main benefit of these statistical rate models (e.g., the linear aging reliability model [Vesely, 1987]) is their simplicity and ease of application, we pointed out that these models do not capture the total impact or behavior of complex aging mechanisms such as FAC.

It was found that FAC in the main feedwater piping over 10 years had only a slight impact on the risk at Surry. The reason for this result is that the contribution to the total plant core damage frequency from loss of main feedwater in the Surry IPE is less than 1%. To reiterate, this feasibility study *should not* be construed to represent the magnitude of risk posed by FAC in LWRs. While indirect effects were considered, a rigorous treatment of their impact was not performed. Considering the lack of contribution to total risk from FAC for the piping that was studied, the indirect effects may become more important. Further, note that the risk metric used in this paper was core damage frequency. Other risk measures (large early release frequency, off-site consequences) could be utilized. In addition, monetary arguments were considered to be outside the scope of the feasibility assessment. The financial risk of events such as a FAC-caused rupture of MFW piping may be important and could be quantified using the general methodology described in this paper.

Serious consideration should be given to applying the techniques described in this report to *non-aging* issues. For example, given the availability of a failure model for pump seals (potentially leading to a loss of coolant scenario), the failure model could be incorporated directly into the PRA rather than just imbedding the *results* of the model in a basic event in the PRA. A similar application could be for embedding an off-site power recovery model directly into a PRA model.

8. References

Apostolakis, G. E., 1995. "A Commentary on Model Uncertainty," in: Proceedings of Workshop on Model Uncertainty," A Mosleh, N. Siu, C. Smidts, and C. Lui, Eds., Center for Reliability Engineering, University of Maryland, College Park, MD (also published as Report NUREG/CP-0138, US Nuclear Regulatory Commission, Washington, DC, 1994).

Arkansas Nuclear One Unit 1 1990. "Reactor Shutdown Required by Technical Specification Due to Unisolable Leak in a Pressurizer Nozzle which was Caused by Pure Water Stress Corrosion Cracking," Licensee Event Report 90-021-00, Unit 1, Docket 50-313, Energy Operations.

ASME 1995. *ASME Boiler and Pressure Vessel Code*, Section XI, Appendix C: Evaluation of Flaws in Austenitic Piping, American Society of Mechanical Engineers, New York, pp. 381-394.

Barsom, J. M., and Vecchio, R. S., 1995. "Fatigue of Welded Components," ASME/JSME Pressure Vessel and Piping Conference, Honolulu, Hawaii.

Bosnak, R. J. 1987. "Meeting Minutes of 1/15/87 Technical Panel Discussion of Surry-2 Pipe Failure

Implications," Memorandum to T.P. Speis, U.S. Nuclear Regulatory Commission.

Chexal, B., et al., 1996, *Flow-Accelerated Corrosion in Power Plants*, EPRI TR-106611, Electric Power Research Institute, Palo Alto.

Chexal, B and J. S. Horowitz, 1995. "Chexal-Horowitz Flow-Accelerated Corrosion Model-Parameters and Influences," *Current Perspectives of International Pressure Vessels and Piping Codes and Standards*, PVP-Vol. B, pp. 231-243. American Society of Mechanical Engineers, New York.

Chopra, O. K. 1992a. "Estimation of Mechanical Properties of Cast Stainless Steels During Thermal Aging in LWR Systems," *Proceedings of the U.S. Nuclear Regulatory Commission Nineteenth Water Reactor Safety Information Meeting*, NUREG/CP-0119, Vol.1, pp.151-178.

Chopra, O. K. 1992b. *Long-Term Aging of Cast Duplex Stainless Steels in LWR Systems*, NUREG/CR-4744, Vol. 6, No.2, ANL-92/32.

Cooper, W. E. 1992. "The Initial Scope and Intent of the Section III Fatigue Design Procedures," presented at the *Pressure Vessel Research Council Workshop on Environmental Effects on Fatigue Performance*, Clearwater Beach, Florida.

Foster, J. P., W. H. Bamford, and R. S. Pathania 1995. "Initial Results of Alloy 600 Crack Growth Rate Testing in a PWR Environment," *Proceedings of the Seventh International Symposium on Environmental Degradation of Materials in Nuclear Power Systems - Water Reactors, August 7-10, 1995*, NACE International, pp. 25-40.

Gosselin, S. R., November 1997, "EPRI's New In-Service Pipe Inspection Program," Nuclear News, American Nuclear Society, La Grange Park, Illinois.

Higuchi, M. A., et al. 1996. "A Study on Fatigue Strength Reduction Factor for Small Diameter Socket Welded Pipe Joints," PVP-Vol. 338, *Pressure Vessel and Piping Codes and Standards*, Volume 1, ASME, pp. 11-19.

Kaplan, S. 1983. "On a 'Two-Stage' Bayesian Procedure for Determining Failure Rates from Experiential Data," *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS 102, No. 1. 195-202.

Kastner, W., and E. Riedle 1986. *Empirical Model for Calculation of material losses due to Corrosion erosion*, VGB Kraftwerkstechnik 66, No.12, pp. 1023-1029.

Kastner, W., 1987, *Erosion-Corrosion Experiments and Calculation Model*, EPRI Workshop on Erosion-Corrosion of Carbon Steel Piping: Nuclear and Fossil Plants.

Keck, R. G., and P. Griffith 1987. *Prediction and Mitigation of Erosion-Corrosive Wear in Secondary Piping Systems of Nuclear Power Plants*, NUREG/CR-5007.

Kooistra, L., et al. 1961. "Full Size Pressure Vessel Testing and Its Approach to Fatigue," *Journal of Engineering for Power, Transactions of the American Society of Mechanical Engineers*, 86, 4, pp. 419-28.

- Remy, F. N. and M. Bouchacourt 1992. "Flow-Assisted Corrosion: A Method to Avoid Damage," *Nuclear Engineering and Design*, 133, p. 23-30.
- Sanzo, D. et al., 1994. *Survey and Evaluation of Aging Risk Assessment Methods and Applications*, NUREG/CR-6157, LA-12715-MS.
- Scott, P. M. 1991. "An Analysis of Primary Water Stress Corrosion Cracking in PWR Steam Generators," *Proceedings of the Specialists Meeting on Operating Experience with Steam Generators*, Brussels, Belgium, September, 1991, Paper 5.6.
- Selby, D. L., et al. 1985. *Pressurized Thermal Shock Evaluation of the H. B. Robinson Unit 2 Nuclear Power Plant*, NUREG/CR-4183, ORNL/TM-9567/V1.
- Shack, W. J., and O. Jonas 1988. "Investigation of Erosion-Corrosion in Reactor Piping," presented at the *Workshop on Structural Integrity of Reactor Piping Systems, May 16-19, Tokyo, Japan*, U.S. Nuclear Regulatory Commission and Japanese Ministry of International Trade and Industry.
- Shah, V. N., et al. 1992. "Assessment of Primary Water Stress Corrosion Cracking of PWR Steam Generator Tubes," *Nuclear Engineering and Design*, 134, pp. 199-215.
- Shah, V. N. and P. E. MacDonald, 1993, *Aging and Life Extension of Major Light Water Reactor Components*, Elsevier Science Publishers.
- Shah, V. N., et al. 1998. *Assessment of Pressurized Water Reactor Primary System Leaks (Final Draft)*, NUREG/CR-6582, INEEL/EXT-98-01068.
- Simonen, F. A., et al. 1986a. *VISA-II - A Computer Code for Predicting the Probability of Reactor Pressure Vessel Failure*, NUREG/CR-4486, PNL-5775, U.S. Nuclear Regulatory Commission, Washington, D.C.
- Simonen, F. A., et al. 1986b. *Reactor Pressure Vessel Failure Probability Following Through-Wall Cracks Due to Pressurized Thermal Shock Events*, NUREG/CR-4483, PNL-5727.
- Siu, N. O. and G. E. Apostolakis, 1982. "Probabilistic Models for Cable Tray Fires," *Reliability Engineering*, 3:213-227.
- Siu, N. O. and G. E. Apostolakis, 1985, *On the Quantification of Modeling Uncertainties*, Transactions of 8th International Conference on Structural Mechanics in Reactor Technology, Brussels, Belgium, August 19-23, Paper Vol. M2. 1/5 Systems Reliability of Nuclear Power Plants, pp. 375-378.
- Siu, N. O., et al., 1992, *Bayesian Assessment of Modeling Uncertainties: Application to Fire Risk Assessment*, Analysis and Management of Uncertainty: Theory and Application, Ayyub, B. M., et al., edit, 1992, Elsevier Science Publishers B.V., pp. 351-361.

Smith, C.L., T. A. Thatcher, and J. K. Knudsen, 1998. *Systems Analysis Programs for Hands-On Integrated Reliability Evaluations (SAPHIRE) Basic Training Course and SAPHIRE Basics Workshop Manual*, INEEL/EXT-1998-1136.

Terrell, J. B. 1988. *Fatigue Life Characterization of Smooth and Notched Piping Steel Specimens in 288°C Air Environments*, NUREG/CR-5013, MEA-2232.

USNRC 1987a. "Feedwater Line Break," USNRC Information Notice 86-106, Supplement 1.

USNRC 1992a. "Rapid Flow-Induced Erosion/Corrosion of Feedwater Piping," USNRC Information Notice 92-07.

USNRC 1992b. "Higher than Predicted Erosion/Corrosion in Unisolable Reactor Coolant Pressure Boundary Piping Inside Containment at a Boiling Water Reactor," USNRC Information Notice 92-35.

Vesely, W. E., 1987, *Risk Evaluation of Aging Phenomena: the Linear Aging Reliability Model and Its Extension*, NUREG/CR-4769, USNRC, Washington, D. C.

Virginia Power 1987. *Surry Unit 2 Reactor Trip and Feedwater Pipe Failure Report*, Rev. 0.

Vora, J. P., 1993, *NRC Research Program on Plant Aging: Listing and Summaries of Reports Issued Through September*, NUREG-1377, Rev. 4, U.S. Nuclear Regulatory Commission.

Wesley, D. A., et al. 1990. *Pressure-Dependent Fragilities for Piping Components*, NUREG/CR-5603, EGG-2607, Nuclear Regulatory Commission, pp. 2-1 to 2-15.

ACCIDENT SEQUENCE PRECURSOR PROGRAM LARGE EARLY RELEASE FREQUENCY MODEL DEVELOPMENT

Douglas A. Brownson
Idaho National Engineering and
Environmental Laboratory
P.O. Box 1625, MS 3850
Idaho Falls, ID 83415-3850
(208) 526-9460
E-mail: dov@inel.gov

Thomas D. Brown
Sandia National Laboratories
P.O. Box 5800, MS 0405
Albuquerque, NM 87185-0405
(505) 844-6134
E-mail: tdbrown@sandia.gov

Felicia A. Duran
Sandia National Laboratories
P.O. Box 5800, MS 0747
Albuquerque, NM 87185-0747
(505) 844-4495
E-mail: faduran@sandia.gov

Julie J. Gregory
Sandia National Laboratories
P.O. Box 5800, MS 0748
Albuquerque, NM 87185-0748
(505) 844-7539
E-mail: jjgrego@sandia.gov

Edward G. Rodrick
U.S. Nuclear Regulatory Commission
Mail Stop T10E50
Washington, DC 20555
(301) 415-5871
E-mail: egr@nrc.gov

ABSTRACT

The U.S. Nuclear Regulatory Commission's (NRC) Accident Sequence Precursor (ASP) program was initiated by the Office of Nuclear Regulatory Research (RES) to provide a probabilistic method for reviewing operational experience to determine and assess both known and previously unrecognized vulnerabilities that could lead to core damage accidents. The ASP program is currently managed by the Office of Analysis and Evaluation of Operational Data (AEOD). Standardized Level 1 plant models developed using the SAPHIRE computer code¹ are used to assess the conditional core damage probability for operational events during full power operation occurring in commercial nuclear power plants (NPPs). Because not all accident sequences leading to core damage will result in the same radiological consequences, work was begun in 1995 to study the feasibility of developing simplified Level 2/3 models using the SAPHIRE computer code to estimate the radiological consequences associated with the radioactive release to the environment resulting from the Level 1 core damage scenarios. Once developed, these simplified Level 2/3 models were linked to the Level 1 models to provide risk perspectives for operational events. Ten Level 2/3 prototype models were completed to represent the various pressurized water reactor (PWR) and boiling water reactor (BWR) nuclear steam supply systems and containment types.

During the development of the simplified Level 2/3 models, several limitations were identified with the prototype models. These limitations were primarily the result of the simplification process necessary to develop these models. In addition, in 1997 the NRC began moving towards using large early release frequency (LERF) as a surrogate for the early fatality quantitative health objectives [refer to Regulatory Guide (RG) 1.174² (draft issued as DG-1061)]. As defined in RG 1.174, LERF is "the frequency of those accidents leading to significant, unmitigated releases from containment in a time frame prior to effective evacuation of the close-in population such that there is a potential for early health effects." These accidents include unscrubbed releases resulting from failure of containment isolation, containment bypass events, and scenarios which result in containment failure at or shortly after reactor vessel breach. The focus of the Level 2/3 ASP project was then modified to address LERF and to correct the limitations identified with the

simplified Level 2/3 models. The development of LERF models is also thought to be a more useful and practical approach for use in event assessment and the AEOD ASP program.

In the spring of 1998, work was begun to develop detailed LERF models for the ten plant classes established during the Level 2/3 model development work. The LERF models, as well as the previously developed Level 2/3 prototype models, are based on the accident progression event trees (APETs) developed during the NUREG-1150 studies.³ Because these models will be "detailed" rather than "simplified", it is envisioned that the initial ten plant models will be capable of being modified to represent any NPP within each plant containment class given sufficient plant-specific data. Modification of the simplified Level 2/3 models to represent other NPPs could not be readily performed, which provided additional impetus to move toward detailed LERF models.

DEVELOPMENT OF ASP LEVEL 2/3 PROTOTYPE MODELS

The ASP LERF model development builds on the work that had been performed during the development of the ASP Level 2/3 prototype models. The ASP Level 2/3 work was initiated in 1995 with the objectives of (1) demonstrating the process for development of simplified Level 2/3 models, (2) determining the information required for the development of these simplified models, (3) developing the appropriate interface between the ASP Level 1 and Level 2/3 models, and (4) integrating the Level 2/3 models into the existing ASP software, SAPHIRE. The work resulted in the development of ten prototype, simplified Level 2/3 models. The NPPs modeled were determined by first sorting all commercial nuclear power plants (NPPs) into groups based on the combination of different containment and nuclear steam supply system designs. This categorization produced four boiling water reactor (BWR) groups and six pressurized water reactor (PWR) groups. These ten prototype models were developed for both NPPs which had detailed probabilistic risk assessments (PRAs) performed during the NUREG-1150 studies and NPPs that did not. Therefore, the process necessary to develop simplified Level 2/3 models when a detailed PRA model was not available was also investigated. The current NPP groups and the prototype NPP selected for each group are listed in Table 1.

The development of the ASP simplified Level 2/3 prototype models required the binning of all ASP Level 1 accident sequences leading to core damage into a plant damage state (PDS) based on the conditions of the reactor coolant system (RCS) and the status of important accident mitigation systems at the time of core damage. Since the end states of the existing ASP Level 1 models indicate only the core status (OK or Core Damage) for the accident sequences, and do not include sufficient containment system information that is crucial for the Level 2 analysis, it was necessary to develop bridge event trees (BETs) which model the required containment systems. The BETs were then linked to those accident sequences leading to core damage. The use of BETs allowed system dependencies between the various systems, such as electrical power, to be accounted for in the fault trees that support the BET top events.

The ASP Level 2/3 modeling process follows the NUREG-1150 Level 2/3 modeling process closely in that detailed accident progression analyses were performed for each of the different PDSs. The results from these detailed analyses were then "rolled-up" to quantify the nine to ten top events of the simplified Level 2 SAPHIRE model. Source terms were estimated for each Level 2 sequence and consequence calculations were performed for either individual source terms or source term groups. These results were then incorporated into a Level 3 SAPHIRE model. Risk results were produced by linking all the analyses.

Table 1. Reactor Groups and Reference NPP for Each Reactor Group

| Reactor Group | Prototype NPP |
|---|----------------|
| Westinghouse PWRs with a large, dry containment | Byron* |
| Combustion Engineering PWRs with PORVs and a large, dry containment | Calvert Cliffs |
| Westinghouse PWRs with a subatmospheric containment | Surry |
| Westinghouse PWRs with an ice condenser containment | Sequoyah |
| Combustion Engineering PWRs without PORVs and with a large, dry containment | Palo Verde |
| Babcock & Wilcox PWRs with a large, dry containment | Oconee |
| BWRs with a Mark I containment (BWR/4 with RCIC) | Peach Bottom |
| BWRs with a Mark I containment (BWR/3 with isolation condenser) | Dresden* |
| BWRs with a Mark II containment | LaSalle |
| BWRs with a Mark III containment | Grand Gulf |

- * For the ASP LERF model development, the Byron NPP replaces the Zion NPP as the prototype NPP for the "Westinghouse PWR with a large, dry containment" reactor group and the Dresden NPP replaces the Brunswick NPP as the "BWR with Mark I containment (reinforced concrete drywell and reinforced concrete with steel liner wetwell)" reactor group.

Although the ASP simplified Level 2/3 prototype models accurately reproduced the detailed analyses, several limitations with these models were identified when the reasonableness of extrapolating the prototype model results to the remaining NPPs in the same reactor group was examined. Because the NPPs of each reactor group differ greatly (physical attributes, equipment employed, and operational procedures), it is unlikely that a Level 2/3 model for one NPP can be used as a surrogate for another or that its results could be extrapolated to the entire reactor group. The generation of a new NPP model would necessitate going back to the supporting detailed model to make the plant specific modifications, analyzing the detailed models with the modifications, and then re-rolling up the supporting detailed model results to create the new simplified model. In addition, each simplified model is created using a predefined set of PDSs. When new PDSs are created as a result of changes to the Level 1 models, the detailed model must be modified and re-analyzed to address these new PDSs. Also, because so many details are hidden in the rolling up process, it was recommended that only the supporting detailed models be modified when performing system importance studies.

In summary, an ASP simplified Level 2/3 prototype model is a simplified representation of an underlying detailed model. The formats and application tools are not compatible, and so translation of the detailed to the simplified model is required. If any significant model changes are necessary, they are implemented in the detailed model whose results are then translated accordingly to the simplified model. This process becomes problematic if many changes are warranted. Therefore, this prototype method contains limitations that are potentially unmanageable if this process is to be extrapolated to other NPPs beyond the prototype

models. In addition, the process required to modify and extrapolate these models would be very resource intensive.

Coupled with the ASP simplified Level 2/3 prototype model limitations is a move by the NRC towards increased risk-informed decisions through the use of a calculated large early release frequency (LERF). Together, these two factors resulted in a refocus of this project to develop a methodology for quantifying the LERF which also eliminates the identified ASP simplified Level 2/3 prototype model limitations.

ASP LERF MODEL DEVELOPMENT

The structure of the ASP LERF models differ from the ASP Level 2/3 prototype models in that the Level 2 accident progression is only developed as far as it takes to determine whether or not a large, early release has occurred. Because it is necessary to develop accident progression sequences only as far as early containment failure, it then became feasible to create models similar to the NUREG-1150 APET models. These models ask a series of questions which query the status of important accident mitigation systems and determine the most likely outcomes of phenomenological events based on the RCS and containment conditions of each accident progression sequence. The main advantages of developing these questions in a SAPHIRE model (rather than analyzing a detailed model outside of SAPHIRE and rolling up the results into a few key event tree top events as was done in the development of the simplified Level 2/3 models), is that the impact of individual systems or phenomena on the overall LERF results can now be readily obtained. Also, modifications in the ASP Level 1 models that may result in the production of new PDSs that were not previously analyzed can be processed without modification to the LERF model. Most importantly, the feasibility of developing the detailed LERF models was made possible by the modifications made to the SAPHIRE computer code which significantly accelerated the processing time of large event trees and the development of faster, more powerful personal computers since this program was initiated in 1995.

The assessment of the LERF is implemented in the ASP LERF models as the frequency of any sequence in which core damage leads to vessel failure and the containment function is compromised. For the PWRs, the containment function is considered compromised by containment bypass events such as failure to isolate, steam generator tube rupture or interfacing systems loss-of-coolant accident (LOCA), or compromised by mechanical failure from phenomena such as over pressurization, impulse loading, missile generation or direct core contact with the containment wall. For the BWRs the containment function is considered compromised if the early releases are not scrubbed by the suppression pool and an event occurs that breaches the primary containment boundary such as: venting, failure to isolate, interfacing systems LOCA, or mechanical failure from phenomena such as over pressurization, impulse loading, missile generation, or direct core contact with the containment wall.

The structure of the ASP LERF models closely mimics that of the APETs developed during the NUREG-1150 studies. Supporting detailed models were developed only as far as early containment failure from the NUREG-1150 APET models in order to assist in the development and verification process of the ASP LERF models. The questions in the supporting detailed model are sequentially developed as SAPHIRE top events in a series of event trees that are linked. These linked trees are referred to as subtrees. The ASP LERF event trees contain between five and 15 subtrees, depending on the size of the model. Each subtree has between three to five top events so that each subtree can easily be maintained within SAPHIRE. The complete APET models developed for the NUREG-1150 studies vary in size from 71 questions (or top

events) for the Surry model to 145 questions for the Peach Bottom model. Development of the supporting detailed models and the LERF models would require modeling 43 and 109 of these questions, respectively.

In the ASP LERF models, when a branch of an accident progression question is encountered whose result can be interpreted as being a contributor to LERF or no LERF (e.g., containment isolation failure and no vessel breach, respectively), further development of that branch is terminated and its frequency binned accordingly to its LERF or no LERF end state. Quantification of the branches that are not terminated occurs through the assignment of split fractions by a series of IF-THEN-ELSE rules. These rules are written for each event tree top event and are based on the logic contained in the NUREG-1150 APETs.

The values used in the quantification of the event tree branches are obtained from three sources, all of which are utilized in the NUREG-1150 APET methodology. The first source is the split fractions applied directly to the APET logic. In this case, the split fractions obtained from the APET logic are applied directly to the ASP LERF model. The second source is a distribution that is randomly sampled. In this case, the average of the randomly sampled distributions is calculated and used as the split fractions in the ASP LERF models. The third source for the quantification of event tree branches are user functions that calculate split fractions based on the input of numerous parameters. These input parameters may include containment pressure at the time of vessel breach and the percentage of zirconium that is oxidized during core degradation (which determines amount of hydrogen produced). In these cases, a matrix of possible input parameter combinations are analyzed by the user functions (each combination referred to as a case) and the resulting split fractions calculated for each case are incorporated into the ASP LERF model as basic events.

The size of the ASP LERF models are smaller than they would be if full Level 2 models were developed; however, concern was expressed over the number of sequences that would be generated by these models and the time that would be required to process them in SAPHIRE. This is especially true for the larger BWR and ice condenser containment PWR models. To address these issues, a number of modeling simplifications were employed to reduce the complexity of the models and therefore the number of sequences generated. In addition, a number of code modifications have been proposed and implemented, greatly accelerating the analysis time of SAPHIRE.

Development of ASP LERF models for NPPs that were not part of the NUREG-1150 studies involves (1) choosing a supporting detailed model that resembles the NPP of interest as closely as possible and (2) making the appropriate plant-specific modifications to that model in order to accurately calculate the split fractions that will be used to represent the NPP of interest. Any necessary plant-specific modifications are then made to the equivalent ASP LERF model and the updated split fractions are then applied. Verification of these models are then performed as outlined below.

VERIFICATION OF LERF MODELS

To date, four of the ten prototype ASP LERF models have been completed. These include two BWR and two PWR models. These models have been verified to ensure that they accurately reproduce the supporting detailed models from which they were developed and produce realistic results.

To ensure that the ASP models accurately reproduce the supporting detailed models, a comparison of the calculated large early release frequencies of the two models was performed for each PDS. Overall, these comparisons resulted in very good agreement. For the LaSalle model, most of the PDS LERF comparisons

resulted in absolute percent differences no greater than 5.0% with absolute difference in the LERF values of no more than 0.05. For the Surry model, the PDS LERF comparisons, the percent differences were calculated to be no more than 0.05%. A sampling of these comparisons is presented for the ATWS PDSs in Table 2. The larger

Table 2. Comparison of ASP LERF Model Results to Supporting Detailed Model Results

| | ASP LERF Model | Supporting Detailed Model | Percent Difference ¹ |
|---|----------------|---------------------------|---------------------------------|
| LaSalle Conditional Large Early Release Probability² Results for ATWS Plant Damage States | | | |
| PDS-ATWS-1 | 9.578E-1 | 9.456E-1 | +1.27 |
| PDS-ATWS-2 | 9.484E-1 | 9.337E-1 | +1.55 |
| PDS-ATWS-3 | 9.484E-1 | 9.337E-1 | +1.55 |
| PDS-ATWS-4 | 9.331E-1 | 9.200E-1 | +1.41 |
| PDS-ATWS-5 | 9.578E-1 | 9.456E-1 | +1.27 |
| Surry Conditional Large Early Release Probability² Results for ATWS Plant Damage States | | | |
| PDS-ATWS-1 | 1.336E-2 | 1.336E-2 | -0.037 |
| PDS-ATWS-2 | 8.397E-3 | 8.397E-3 | -0.002 |
| PDS-ATWS-3 | 1.423E-2 | 1.423E-2 | +0.021 |
| PDS-ATWS-4 | 1.146E-2 | 1.146E-2 | -0.009 |
| PDS-ATWS-5 | 1.146E-2 | 1.146E-2 | -0.009 |

100

1. Percent Difference = [(Supporting Model LER Prob. - ASP Model LER Prob.)/Supporting Model LER Prob.] X

2. Conditional LER probabilities are calculated by setting the individual initiating event PDS frequencies to 1.00.

percent differences of the LaSalle model can be attributed to the number and type of simplifications that were made to this model in order to reduce its size for incorporation into SAPHIRE. In addition, the process that is used to quantify some SAPHIRE model split fractions from sampled distributions does not necessarily reproduce the realized split fractions in the supporting detailed model because of the unavoidable differences in the sampling process.

An additional verification of the ASP LERF models was performed by comparing the calculated ASP LERF results to equivalent results from a full scope PRA study. The results of the comparison of the Surry ASP LERF models results to the Surry NUREG-1150 study⁴ are presented in Table 3. As shown in this table, the magnitude of the calculated ASP core damage frequency (CDF) and total LERF results are below those obtained from the NUREG-1150 study. The differences in the LERF results for the two models arises mainly from the differences in CDF (a comparison of the conditional LERF probabilities given core damage

indicates relatively good agreement between the ASP LERF model and the NUREG-1150 study results). The difference in CDF results are due to differences in initiating events modeled in the ASP Level 1, Rev. 2 QA (quality assured) models versus the NUREG-1150 study. Only those initiating events which have the highest frequency are currently being modeled in the ASP Level 1, Rev. 2 QA models; whereas, more initiating events were modeled in the NUREG-1150 study. The initiating events modeled in the ASP Level 1 analysis include loss-of-offsite power, steam generator tube rupture (SGTR), small break loss-off-coolant accidents, and transients. Future work scope

Table 3. Comparison of Surry ASP LERF Model end state results to equivalent Surry NUREG-1150 study results

| | Frequency (reactor-year ⁻¹) | Conditional Probability * |
|---|---|---------------------------|
| Total Core Damage Frequency | | |
| ASP LERF Model | 3.2E-5 | N/A |
| NUREG-1150 Model | 4.1E-5 | N/A |
| LERF Resulting from Induced Steam Generator Tube Rupture Events | | |
| ASP LERF Model | 2.9E-6 | 9.1E-2 |
| NUREG-1150 Model | 1.9E-6 | 8.3E-2 |
| LERF Resulting from Bypass Events (no containment isolation, SGTR initiating events) | | |
| ASP LERF Model | 1.2E-8 | 3.6E-4 |
| NUREG-1150 Model | 6.4E-9 | 1.5E-4 |
| LERF Resulting from Early Containment Failure Events (vessel breach, Alpha mode, rocket) | | |
| ASP LERF Model | 2.9E-7 | 9.0E-3 |
| NUREG-1150 Model | 2.7E-7 | 7.2E-3 |
| Total LERF | | |
| ASP LERF Model | 3.2E-6 | 1.0E-1 |
| NUREG-1150 Model | 3.8E-6 | 1.3E-1 |

* Conditional probability calculated by dividing the LERF frequency by the total plant damage state (PDS) frequency
 N/A - not applicable

of the ASP Level 1 models includes the development of additional initiating events (i.e. large break, medium break, and interfacing system LOCAs) which will capture a greater percentage of the overall CDF and subsequently capture a greater percentage of overall LERF for any given NPP.

VERIFICATION OF ASP LERF METHODOLOGY

The definition of LERF used for the development of the ASP LERF models is consistent with the definition stated in RG 1.174. However, RG 1.174 does not directly present an approach for determining the LERF, but instead cites a simple screening approach for calculating the LERF as documented in NUREG/CR-6595⁵. Additional supporting information for NUREG/CR-6595 and RG 1.174 is presented in References 6 and 7. The ASP LERF methodology is in general agreement with the intent of the NUREG/CR-6595 approach, however, there are a few exceptions in implementation. The most prominent difference is that the ASP LERF models are much more detailed than the models described in NUREG/CR-6595. The NUREG/CR-6595 models were designed to be simplified models with more general application. The simplified general nature of a NUREG/CR-6595 model creates a problem when trying to investigate the precursor event implications for a containment system. In other words, the simplifications that are adopted in the NUREG/CR-6595 models do not accommodate the analysis of a precursor event. Thus, more detailed LERF models were adopted for application in the ASP program.

In essence, the less complex NUREG/CR-6595 models are included in the LERF models, providing consistency between the two methodologies. During the development of the ASP LERF models, a top event was added that addresses the timing of the release and the relative potential for evacuation of the close-in population (see Figure 1). This query is necessary to fulfill the latter part of the definition of LERF as given in RG 1.174 - "*the frequency of those accidents leading to significant, unmitigated releases from containment in a time frame prior to effective evacuation of the close-in population such that there is a potential for early health effects.*" (In the development of the ASP LERF models, early health effects are interpreted as early fatalities since, according to RG 1.174, "LERF is being used as a surrogate for the early fatality QHO [quantitative health objective].") This feature was not included in the NUREG-1150 APET models because the issue was addressed in the source term and consequence analysis portions of the PRA.

To effectively quantify the "No Early Fatalities" top event shown in Figure 1, it is necessary to determine a source term for each of these sequences and then calculate consequence results for each source term. The consequence calculations would require site-specific knowledge of the circumstances under which a general emergency would be declared, how soon after the accident evacuation would begin, evacuation speed, and meteorological conditions. It was beyond the scope of this project to determine these factors. Therefore, all LERF accident progression sequences are assumed to result in early fatalities. However, it is possible with these models to make exceptions for those sequences whose duration from accident initiation to radiological release is so long that it is unreasonable to assume that an effective evacuation of the close-in population could not occur in a time frame that would prevent early fatalities. An example of such a sequence would be a long-term station blackout scenario in which the duration between core damage and vessel failure could be as long as sixteen hours. The "NO EARLY FATALITIES" top event has been included in these models to allow quantification of these sequences at a later date when sufficient information becomes available.

| Initiating Event | Early Containment Failure | No Early Fatalities | End State |
|------------------|---------------------------|---------------------|-----------|
| IE | ECF | NEF | |
| | Containment Failure | No Early Fatalities | NOLERF |
| | | Early Fatalities | LERF |
| | No Containment Failure | | NOLERF |

Figure 1: Example of implementation of “NO EARLY FATALITIES” top event in ASP LERF model event trees

Because the ASP LERF methodology does not explicitly follow the LERF methodology of NUREG/CR-6595, there is a legitimate concern that ASP LERF results may at times be overly conservative (i.e., over predict the LERF). On the other hand, the models of NUREG/CR-6595 are intended to be simple screening method and may at times produce more conservative results. This is important because RG 1.174 proposes an annual average LERF acceptance guideline of $1E-5$ per reactor year and for plant modifications the change in LERF cannot exceed the acceptance guideline of $1E-6$ per reactor year. Overly conservative results may result in an NPP exceeding these acceptance guidelines.

Reference 6 documents case studies in which the five containment type LERF models of NUREG/CR-6595 were implemented for five U.S. nuclear plants (Surry, Sequoyah, Peach Bottom, Limerick and Grand Gulf). Results of these case studies were compared with the IPE Level 2 results for these plants. In order to compare the IPE Level 2 results to the LERF, three reporting methods were used: (1) frequency of early containment failure or bypass; (2) frequency of a source term release fraction > 0.03 of I and Cs; and (3) frequency of a source term release fraction > 0.1 of I and Cs. The results for the two PWR plants Surry and Sequoyah indicate agreement between the LERF and IPE results, with the LERF estimates slightly higher than the IPE results. The comparison for the Peach Bottom and Limerick BWR plants (Mark I and II containments, respectively), show similar results to those for the PWR plants; however, for these two plants, there is slightly less agreement between the LERF estimates and the I and Cs release fraction frequencies. The Grand Gulf (Mark III BWR containment) case study indicates lower LERF estimates for two of the three IPE methods of comparison. The low estimates of LERF for the Grand Gulf case study were mainly attributed to the exclusion of such things as venting strategies that were included in the IPE but not in the NUREG/CR-6595 LERF model.

Insights provided from the case studies in Reference 6 indicate that large early releases for the PWR plants are dominated by bypass events, and the NUREG/CR-6595 models seem to be appropriate, if not slightly conservative, for estimating LERF. An issue of concern discussed for the PWRs is associated with

the determination of RCS depressurization, including the importance of human error analysis as well as incorporating information from deterministic analyses and the possibility of induced failures. Insights are also noted for the application of the NUREG/CR-6595 BWR LERF models for a few important issues: the LERF models applied somewhat high estimates (compared to the IPE) for the probability of Mark I liner melt-through, venting strategies should be explicitly included in the LERF models, and although suppression pool scrubbing is addressed in the LERF models, there is significant source term mitigation when releases are scrubbed early even if they later bypass the pool. The case study issues discussed in Reference 6 that span both PWR and BWR LERF methods include the determination of igniter availability when fault trees are not included and the concern that although the NUREG/CR-6595 LERF methodology allows the potential to evacuate for long-term sequences (such as loss of containment heat removal), the application of such an assumption may have to be verified and may be non-conservative.

The concerns described in Reference 6 are addressed in the ASP LERF models. The PWR ASP LERF models incorporate the possibility of operator error in the consideration of RCS depressurization; additionally, the assessment of RCS depressurization incorporates information from deterministic analyses including the potential for induced failures in the RCS pressure boundary that are not in the vessel. The BWR ASP LERF models include venting strategies (both containment and MSIV), and the Mark I liner melt-through probabilities from NUREG-1150 are adopted (the inclusion of more recent information is being considered). If the suppression pool is never bypassed through the events that accompany vessel failure, LERF does not occur even if the containment fails (failure would have to be in the wetwell of a Mark I or Mark II for this to happen). As noted in Reference 6, this approach can be overly conservative if earlier releases are scrubbed and the pool is later bypassed; however, the method was intended to credit scrubbing only if all the early releases are predicted to be scrubbed. Hydrogen ignition systems for ice condensers and Mark IIIs are included in the LERF models. The potential to evacuate for long term sequences is included in the ASP LERF models, but the quantification of this capability is currently under development. Overall, the ASP LERF models contain sufficient detail to capture important phenomenological events that a simplified model cannot incorporate.

SUMMARY AND CONCLUSIONS

The objectives for the ASP LERF model development work is to build a Level 2 containment response model that would capture all of the events necessary to define LERF as outlined in RG 1.174, can be directly interfaced with the existing Level 1 models, is technically correct, can be readily modified to incorporate new information or to represent another plant, and can be executed in SAPHIRE.

The ASP LERF models being developed will meet these objectives while providing the NRC with the capability to independently assess the risk impact of plant-specific changes proposed by the utilities that change the NPPs licensing basis. Together with the ASP Level 1 models, the ASP LERF models provide the NRC with the capability of performing equipment and event assessments to determine their impact on a plant's LERF for internal events during power operation. In addition, the ASP LERF models are capable of being updated to reflect changes in information regarding the system operations and phenomenological events, and of being updated to assess the potential for early fatalities for each LERF sequence. As the ASP Level 1 models evolve to include more analysis capabilities, the LERF models will also be refined to reflect the appropriate level of detail needed to demonstrate the new capabilities.

An approach was formulated for the development of detailed LERF models using the NUREG-1150 APET models as a guide. The modifications to the SAPHIRE computer code have allowed the development of these detailed models and the ability to analyze these models in a reasonable time. Ten reference LERF plant models, including six PWR models and four BWR models, which cover a wide variety of containment and nuclear steam supply systems designs, will be complete in 1999. These reference models will be used as the starting point for developing the LERF models for the remaining NPPs.

REFERENCES

1. K. Russell, et al., *Systems Analysis Programs for Hands-on Integrated Reliability Evaluations (SAPHIRE) Version 5.0*, NUREG/CR-6116, EGG-2716, Idaho National Engineering Laboratory, Idaho Falls, Idaho, July 1994.
2. U.S. Nuclear Regulatory Commission, *An Approach for Using Probabilistic Risk Assessment in Risk-Informed Decisions on Plant-Specific Changes to the Licensing Basis*, Regulatory Guide 1.174, Washington, DC, June 1997.
3. U.S. Nuclear Regulatory Commission, *Severe Accident Risks: An Assessment for five U.S. Nuclear Power Plants*, NUREG-1150, Washington, DC, December 1990.
4. R.J. Breeding et.al., *Evaluation of Severe Accident Risk: Surry Unit 1*, NUREG/CR-4551, SAND86-1309, Vol. 3, Rev. 1, Parts 1 and 2, Sandia National Laboratories, Albuquerque, NM, October 1990.
5. W. T. Pratt, et al., *An Approach for Estimating the Frequencies of Various Containment Failure Modes and Bypass Events*, Draft NUREG/CR-6595, BNL-NUREG-52539, Brookhaven National Laboratory, Upton, New York, November 1997.
6. T. Chu and M. A. Azarm, *An Evaluation of the Simplified Event Trees Described in Appendix B of Draft Regulatory Guide (DG-1061)*, letter report WTP/dms97L15, Brookhaven National Laboratory, Upton, New York, October 1997.
7. T. Chu and M. A. Azarm, *Extension of Scope of the Simplified Event Trees Described in Appendix B of Draft Regulatory Guide (DG-1061)*, letter report WTP/dms97L16, Brookhaven National Laboratory, Upton, New York, October 1997.



United States Nuclear Regulatory Commission

Perspectives on Needed Scope and Level of Detail for a PRA Standard

26 -- Water Reactor Safety Meeting -- 1998

Authors:

US Nuclear Regulatory Commission:

Mary Drouin

Gareth Parry

John Lane

Mike Cheok

Lee Abramson

Steve Long

Science Applications International Corp.:

Jeff LaChance

Alan Kolaczowski

Buttonwood Consulting:

Dennis Bley

Sandia National Laboratories:

Allen Camp

John Forester

Donnie Whitehead

Brookhaven National Laboratory:

John Lehner

Trevor Pratt

Ali Azam

Future Resources Associates:

Robert Budnitz

OBJECTIVES AND USES

Objectives:

- ⇒ Define the scope and attributes of a PRA
- ⇒ Enhance efficiency of developing and reviewing a PRA
- ⇒ Contribute to regulatory stability

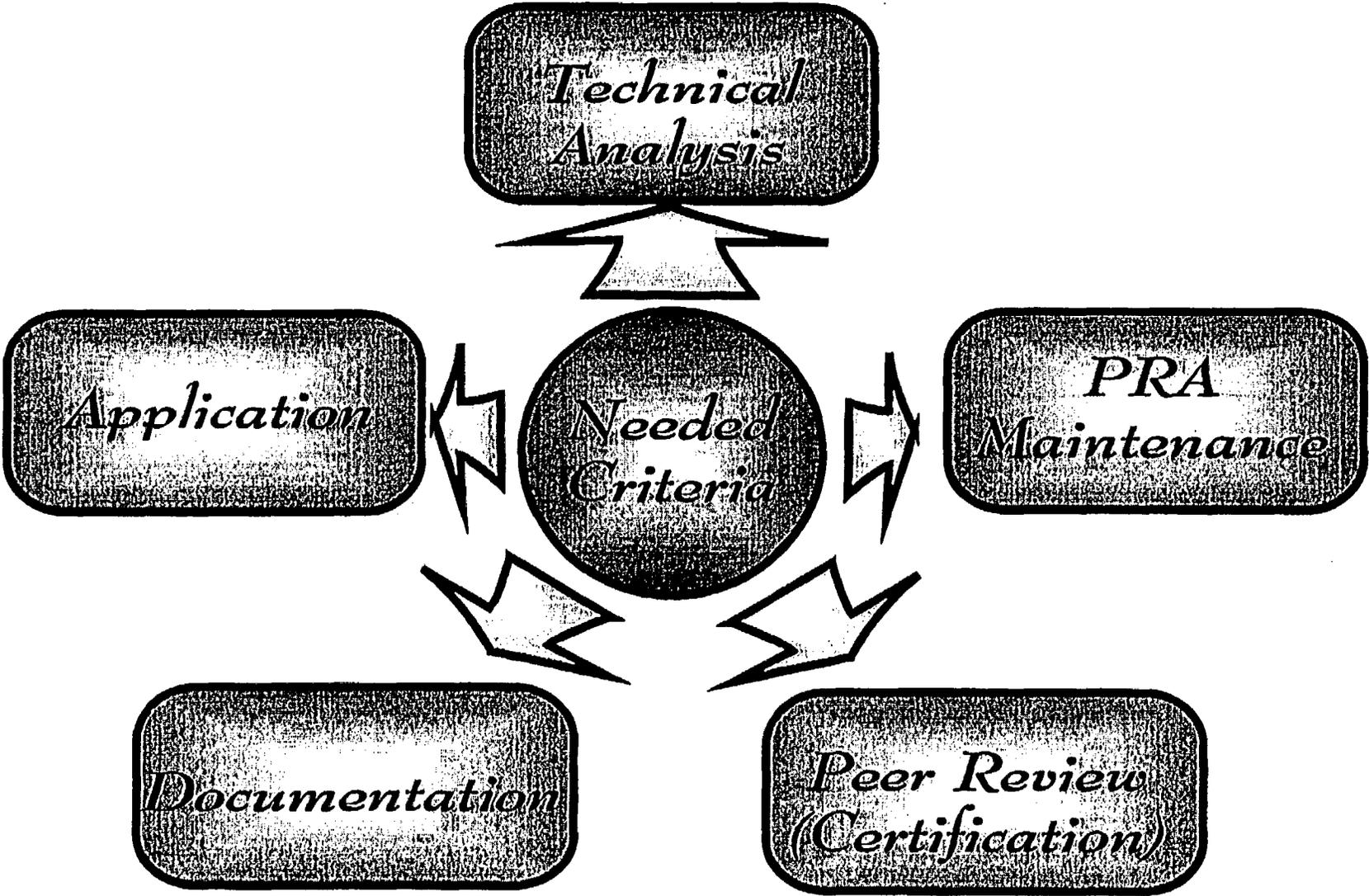
Uses:

- ⇒ Used to support applications to change the licensing basis
- ⇒ NRC to consider endorsing, if successful

PRA STANDARD DEVELOPMENT

- ⇒ ASME leading development of PRA standard
- ⇒ NRC staff working closely with ASME
- ⇒ 1998, scope includes Level 1, internal events (excluding fire), full power operation, and LERF
- ⇒ 1999+, scope includes fire, external events, and low power shutdown

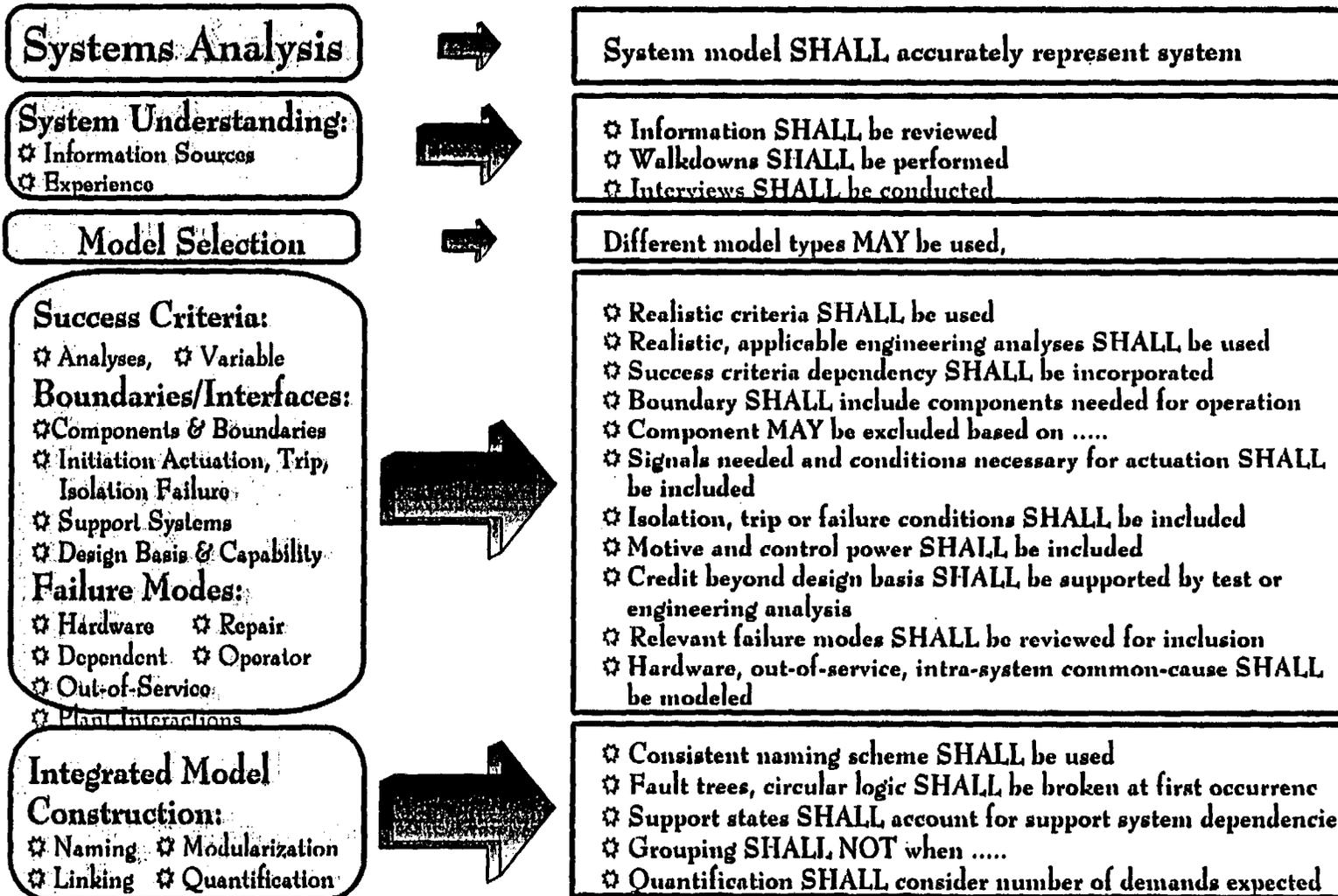
FRAMEWORK



TECHNICAL ANALYSIS REQUIREMENTS

- ⇒ Identification of Essential Elements
- ⇒ Identification of Necessary Functions for Each Element (i.e., "Familiarization, Model Selection, Technical Performance and Integration/Construction)
- ⇒ Identification of Necessary Attributes/Characteristics for Each Function
- ⇒ Elements and Associated Attributes to Produce Realistic Estimation of Core Damage Frequency

TECHNICAL ANALYSIS REQUIREMENTS

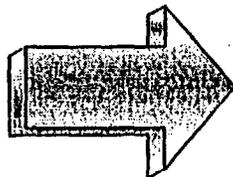


PRA MAINTENANCE REQUIREMENTS

361

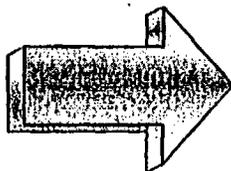
When to Update:

- ✧ Change Driven
- ✧ Time Driven



- ✧ Changes affecting plant SHALL be reviewed
- ✧ PRA SHALL be updated to include effect of change if:
 - 1-- make decisions in the plant, or support licensing application, or used for past decisions; AND
 - 2 -- decisions impacted by the change
- ✧ Plant changes SHALL be reviewed every 2 years
- ✧ Aggregate of changes SHALL be reviewed.....

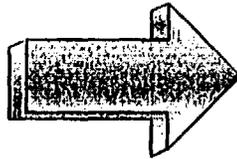
What to Update



- ✧ Change to the plant SHALL be reviewed against requirements
- ✧ For Systems Analysis, the following, at a minimum SHALL be checked to determine if they have been impacted by the changes, and therefore, SHALL be updated: success criteria, ...

PEER REVIEW REQUIREMENTS

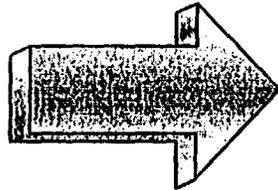
Qualifications & Credentials



- ✧ Review SHALL be conducted by team
- ✧ Team SHALL have collectively necessary expertise
- ✧ Team SHALL be independent of PRA
- ✧ Team leader SHALL be independent of organization
- ✧ Team SHALL have formal PRA training
- ✧ Team SHALL have 10 years of collective PRA experience
- ✧ PRA experience SHALL include 2 different plants
- ✧ Team member SHALL have 5 years in area of responsibility
- ✧ Team member SHALL have plant-specific knowledge
- ✧ Team leader SHALL have plant type knowledge and broad PRA knowledge

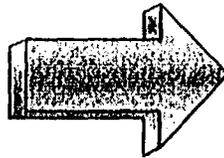
Base PRA Review:

- ✧ Workplan
- ✧ Detailed
- ✧ Limited



- ✧ Team SHALL review workplan
- ✧ Exclusion of requirements SHALL be justified
- ✧ Detailed review SHALL be performed on each system model if no workplan exists
- ✧ Detail review SHALL include examination against each technical requirement
- ✧ Limited review SHALL be performed on each system model not selected for detail review
- ✧ Limited review SHALL include essential elements involving.....

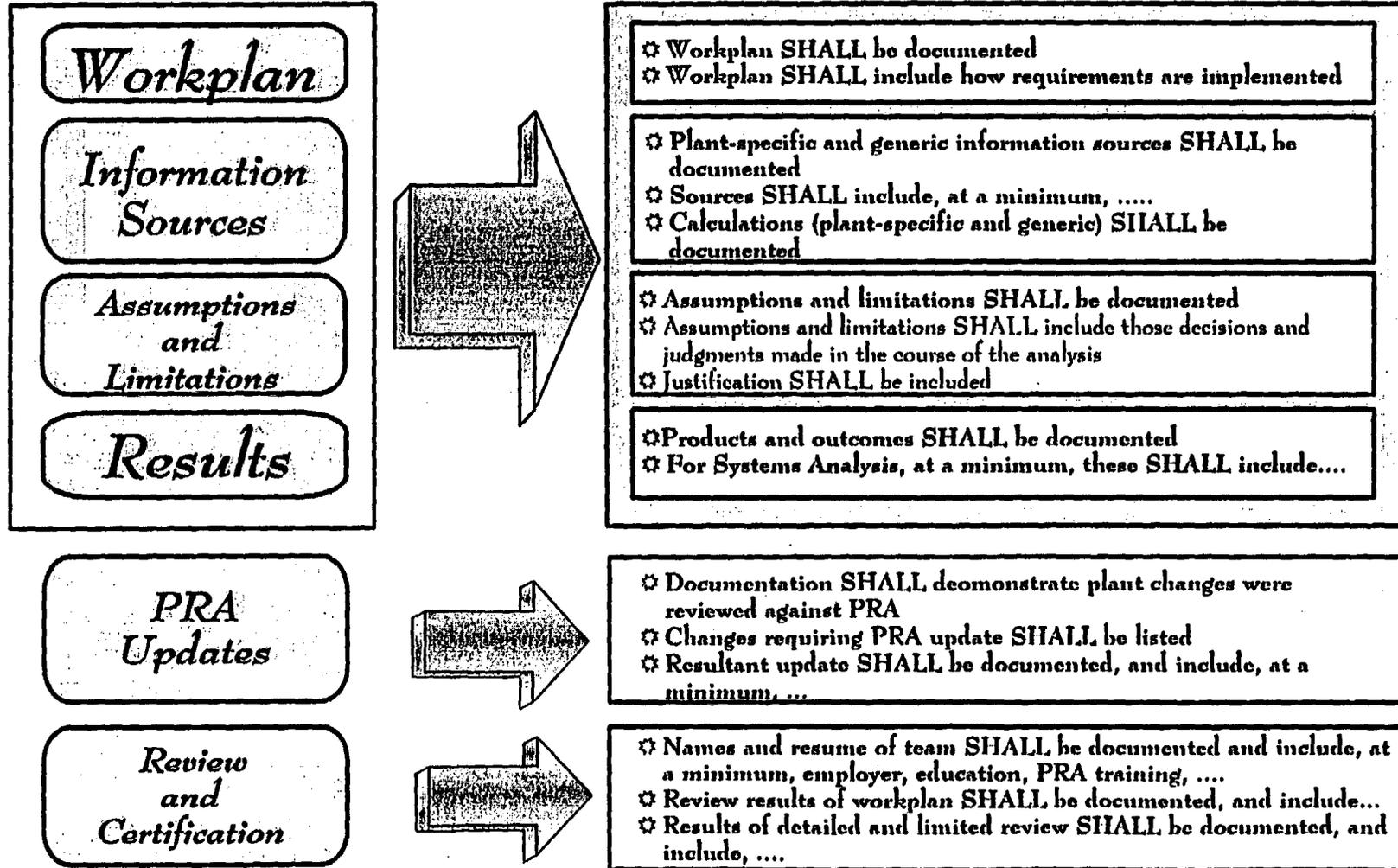
Updated PRA Review



- ✧ Team SHALL review changes made in base PRA to determine requirements accurately implemented
- ✧ Team SHALL use same review requirement as stated for base PRA review

DOCUMENTATION REQUIREMENTS

363



APPLICATIONS

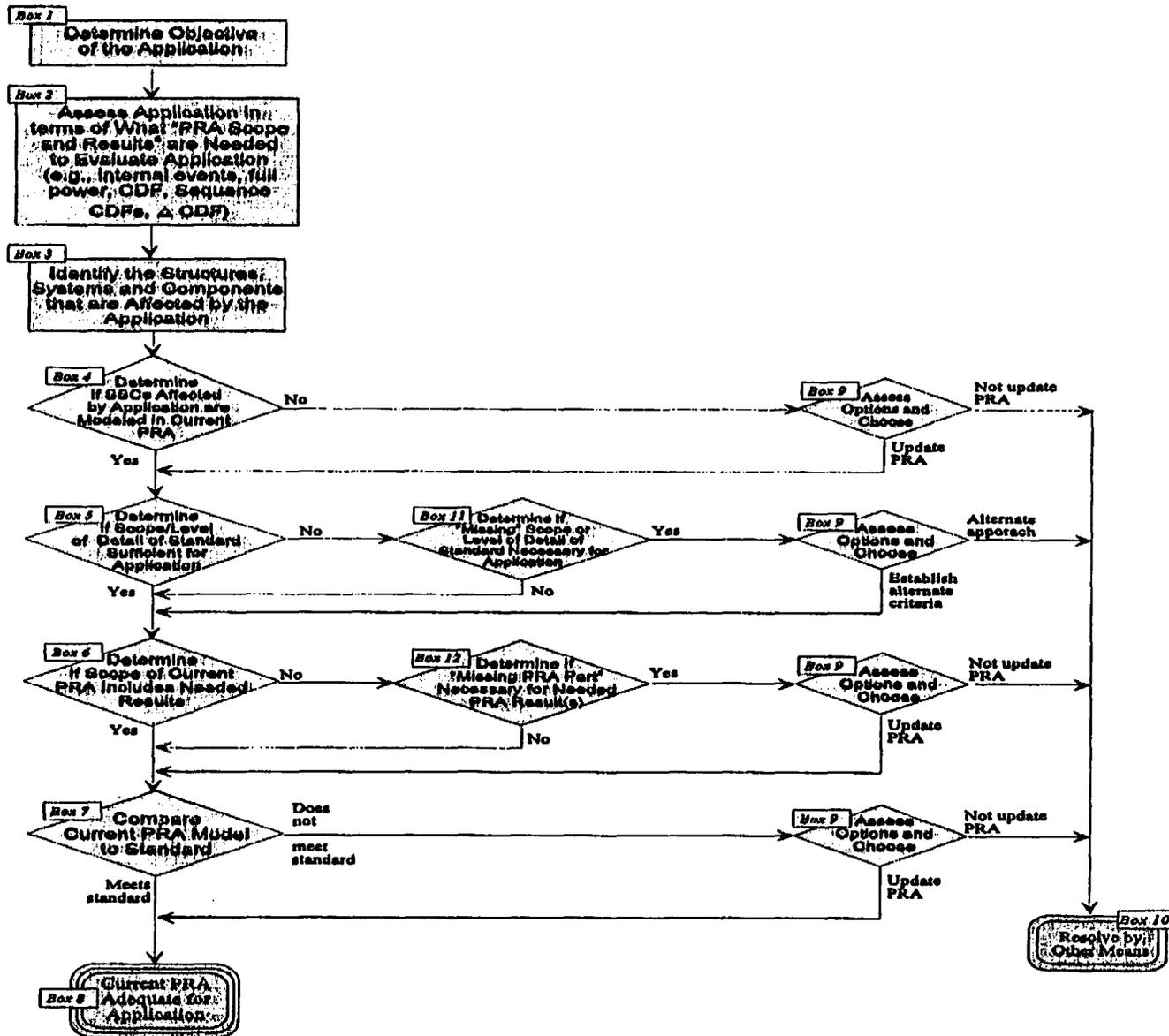
- ✱ **Determine, on application basis, if the PRA, as compared to the standard, is complete**

- ✱ **Determine, on application basis, if standard is "complete"**

- ✱ **Resolution --**
 - ⇨ **Update PRA to the standard**
 - ⇨ **Expand Standard**
 - ⇨ **Resolve by other means (e.g., expert panel)**

APPLICATIONS -- Example Criteria

- ✧ The applicant **shall** determine which elements of the plant, or aspects of its operation are affected by the change.
- ✧ The applicant **shall** determine which PRA elements reflect these plant elements that are affected.
- ✧ The applicant **shall** develop a relationship that characterizes how the PRA elements are changed.
- ✧ The applicant **shall** determine what are the PRA contributors to the change in CDF/LERF.
- ✧ The applicant **shall** compare the treatment of those PRA contributors with the standard and, either, modify the model to meet the standard, OR, provide arguments why not meeting the standard is OK.



NRC SUPPORT FOR THE KALININ (VVER) PROBABILISTIC RISK ASSESSMENT*

**D. Bley
Buttonwood Consulting, Inc.**

**D.J. Diamond, T-L. Chu, A. Azarm, W.T. Pratt
Brookhaven National Laboratory**

**D. Johnson
PLG, Inc.**

**A. Szukiewicz, M. Drouin, A. El-Bassioni, T-M. Su
U.S. Nuclear Regulatory Commission**

Abstract

The U.S. Nuclear Regulatory Commission (NRC) and the Federal Nuclear and Radiation Safety Authority of the Russian Federation have been working together since 1994 to carry out a probabilistic risk assessment (PRA) of a VVER-1000 in the Russian Federation. This was a recognition by both parties that this technology has had a profound effect on the discipline of nuclear reactor safety in the West and that the technology should be transferred to others so that it can be applied to Soviet-designed plants. The NRC provided funds from the Agency for International Development and technical support primarily through Brookhaven National Laboratory and its subcontractors. The latter support was carried out through workshops, by documenting the methodology to be used in a set of guides, and through periodic review of the technical activity. The result of this effort to date includes a set of procedure guides, a draft final report on the Level 1 PRA for internal events (excluding internal fires and floods), and progress reports on the fire, flood, and seismic analysis. It is our belief that the type of assistance provided by the NRC has been instrumental in assuring a quality product and transferring important technology for use by regulators and operators of Soviet-designed reactors. After a thorough review, the report will be finalized, lessons learned will be applied in the regulatory and operational regimes in the Russian Federation, and consideration will be given to supporting a containment analysis in order to complete a simplified Level 2 PRA.

This work was performed under the auspices of the U.S. Nuclear Regulatory Commission.

Origin of the BETA Project

The Kalinin Probabilistic Risk Assessment (PRA) project was designed to improve reactor safety and regulation in the Russian Federation (R.F.), by enhancing the political and technical position of the regulatory agency in Russia and by building a framework and language to address reactor safety issues. The origins of the project lie in the Lisbon Conference on Assistance to the Nuclear Safety Initiative, held in May 1992, where it was agreed that special efforts should be undertaken to improve the safety of the nuclear power plants designed and built by the former Soviet Union. In the following year, the Gore-Chernomyrdin Commission (GCC) was established to improve technical cooperation between the U.S. and the R.F. The U.S. Nuclear Regulatory Commission (NRC) was to provide support to the GCC in nuclear safety, including support to Gosatomnadzor (GAN), the Federal Nuclear and Radiation Safety Authority of the Russian Federation. A November 1993 Memorandum of Meeting (MoM) between NRC and GAN recorded agreement for NRC and GAN to work together, including provision of support to the R.F. to perform a PRA on a VVER-1000 PWR. This was a recognition by both NRC and GAN that this technology has had a profound effect on the discipline of nuclear reactor safety in the West and that the technology should be transferred to others so that it can be applied to Soviet-designed plants. Unit 1 at the Kalinin Nuclear Power Station (KNPS) was chosen for the PRA, and the effort was carried out under the auspices of GAN with the assistance of several other Russian organizations:

- GAN's Science and Engineering Centre for Nuclear and Radiation Safety (SEC-NRS) - the regulatory agency's semi-independent support organization
- Experimental and Design Office "Gidropress" (EDOGP) - the VVER designer
- Nizhny Novgorod Project Institute "Atomenergoprojekt," (NIAEP) - the architect-engineer
- Kalinin Nuclear Power Station (KNPS)
- Rosenergoatom Consortium - the "owner" of KNPS

The MoM addressed how to manage such a project given the many organizations in the R.F. that would need to cooperate to ensure success. A phased approach was to be used with completion of the work in each phase before initiation of the work in a subsequent phase. The four phases of the PRA at Kalinin Nuclear Power Station (later known as the "BETA Project") were to include:

- Phase I. Project Organization
- Phase II. Training, procedure guide development, and information gathering
- Phase III. System modeling and accident frequency analysis (Level 1 PRA; internal and external events)
- Phase IV. Containment performance and risk assessment (simplified Level 2 PRA)

Phase I Activities

During Phase I, a plan was developed for the PRA, including definitions of tasks, levels of effort, schedules, and products. Two primary documents were developed in the U.S. with these specifications:

- General Plan for VVER-1000 Probabilistic Risk Assessment
- Detailed Task Descriptions for the VVER-1000 PRA Project

The tasks planned for the PRA are listed in Table 1.

Table 1. PRA Task List

| Task | Task Title |
|-------------|--|
| III.A | Plant Familiarization and Information Gathering |
| III.B | Identification and Selection of Site Sources of Radioactive Releases |
| III.C | Determination and Selection of Plant Operating States |
| III.D | Definition of Core Damage States or Other Consequences |
| III.E | Selection and Grouping of Initiating Events |
| III.F | Functional Analysis and Systems Success Criteria |
| III.G | Event Sequence Modeling |
| III.H | System Modeling |
| III.I | Human Reliability Analysis |
| III.J | Qualitative Dependence Analysis |
| III.K | Assessment of the Frequency of Initiating Events |
| III.L | Assessment of Component Reliability |
| III.M | Assessment of Common Cause Failure Probabilities |
| III.O | Initial Quantification of Accident Sequences |
| III.P | Final Quantification of Accident Sequences |
| III.R | Interpretation of Results; Importance and Sensitivity Analysis |
| III.S | Spatial Interactions |
| III.T | Fire Analysis |
| III.U | Flood Analysis |
| III.V | Seismic Analysis |
| III.W | Documentation |

The plan for carrying out the PRA was discussed with the Russian team members at a meeting held in May 1995, at GAN in Moscow and at KNPS. The plan was incorporated into formal Implementing Agreements which delineated the responsibilities of NRC and each of the six Russian organizations participating in the project, including funding, schedules, and deliverables.

These first Agreements defined work to be accomplished during the first year of the BETA Project, including Phase II and the initial work on Phase III, the Level 1 PRA. Subsequent meetings were held in Moscow in August 1996 and May 1997 to negotiate Addenda to the Implementing Agreements for each coming year. The NRC would provide financial support for the PRA with funds from the Agency for International Development and technical support primarily through Brookhaven National Laboratory (BNL) and its subcontractors.

Phase II and Early Phase III Work

Phase II was to provide training, to develop procedure guides for the PRA tasks, and to collect information on the plant. The technical work of the project began with a series of workshops. First was a VVER training program for American members of the BETA team. This was held in the R.F. in December 1995. The PRA workshops for the Russian team members consisted of one 8-week long workshop at BNL at the start of the project, followed by 1-week workshops in Russia approximately every six weeks over a period of one and one-half years. The first workshop took place after the initial plant familiarization and information gathering. It consisted of scheduled seminars to provide training on specific technical issues (e.g., development of event sequence diagrams), independent work by the Russian PRA team with interaction with the U.S. experts as needed, and meetings with the U.S. experts to review work in progress. The followup workshops were on technical subjects that enter into the analysis at later times (e.g., human reliability analysis) and subjects that needed further elucidation (e.g., common cause failure analysis).

The procedure guides complemented the workshops. The first draft of the guides used for the Kalinin PRA were prepared in the U.S., reviewed by the R.F. team and translated into Russian. A final version [1] is to be published to be of assistance to other PRA practitioners, especially those with VVER plants. The procedure guides are limited to accidents involving the reactor core and that occur while the plant is operating at full power. Internal initiating events, including internal fires and floods, are considered as well as seismic events. Guidance is provided for a Level 1, 2, and 3 PRA with the Level 3 PRA guidance limited to offsite consequences.

It was assumed that the team carrying out the PRA would be familiar with the set of guides developed by the International Atomic Energy Agency (IAEA) for carrying out a Level 1 PRA for internal events [2]. The IAEA document represented an internationally acceptable approach. The new guides improve on the existing guides by: (1) taking into account recent work in the field, (2) considering special problems that might be specifically present for the VVER experience, and (3) improving upon the guidance already provided. The idea was not to duplicate the existing guidance found in the IAEA document or the material in other guides that have been produced by the NRC [3, 4]. For subjects not well documented in the open

literature (e.g., the approach taken for human reliability analysis), detailed guidance was given; for tasks where a firm understanding was already well established and documentation freely available (e.g., system modeling), minimal guidance and appropriate references were provided.

Phase III, the carrying out of the Level 1 PRA, began with a two-month workshop held at BNL February-April 1996. A series of seminars on specific PRA tasks were held during the visit to BNL. During this intensive training, great strides in the analysis were accomplished. Unfortunately, after the BETA team's return to Russia, progress continued at a slower pace because the team is widely disbursed and involved in other work.

The product requirements for the Phase III PRA include:

- Databases: component failure rates for all VVER-1000 plants and a KNPS-specific database
- IRRAS computer model representing the KNPS Level 1 PRA
- Documentation on the analysis and results
- RELAP5 model for KNPS Unit 1 with all important plant systems
- Level 1 PRA for the KNPS

A Technical Review Group (TRG) was set up with U.S. experts and Russian team members. TRG sessions were held periodically in Moscow (usually in conjunction with training workshops) to review Russian progress on the analysis. The reports of these meetings provided guidance to the team for continuation of the work.

By fall 1997, all Level 1 internal event tasks (excluding fire and flood events) had been completed except final quantification, uncertainty and sensitivity analysis, and final documentation. A number of problems were identified, and an approach and budget for completing the Level 1 PRA were agreed upon.

An additional task was added in 1997 to come up with an Applications Plan. This plan would include a description of how the results/insights of the PRA will be implemented/disseminated at KNPS and other VVER plants. It was to describe those areas where PRA information can be used to improve the plant operations and safety, how this information will be used, and how it will be communicated to other VVER-1000 plants.

Current Status of Phase III

The draft report of the internal events PRA (excluding fire and flood analysis) was submitted in June 1998, and progress reports on the fire, flood, and seismic analysis were also submitted during the year. Unfortunately, during 1998, funding problems precluded having U.S. experts do a detailed review. A cursory review, however, of the internal events report has recently been completed. That brief review has identified a number of potential problems remaining in the PRA and its documentation and many questions that can only be answered by a detailed review and by discussions with the BETA team.

The most important issues that need to be resolved are whether sump plugging is as important as suggested by the results and whether the analysis of scenarios that were thought to be important in the past, but have now been downgraded, is valid. Of the many specific technical questions that are important is the question of why the diesel generator failure rate is unusually low. While it is possible to cite additional examples of items that are questionable beyond those already identified, it may be unfair to do so as many of the questions may be easily answered and determined not to be significant if discussed with the Russian team.

It is also important to note that the English is difficult to read in many places and much of the text is abbreviated and insufficient to explain to a reviewer what was being done. Although a different style of writing is not unexpected when dealing with a different culture, this is a subject that the U.S. experts have repeatedly emphasized in many discussions with the Russian BETA team.

Concluding Remarks

The efforts described in this paper have brought the completion of the PRA for internal events within sight. It is expected that early in 1999 the review of existing documents will be completed, and the final Level 1 PRA report will be made public. Ongoing work is being done to add results for internal fires and floods and seismic events. It is our belief that the type of assistance provided by the NRC has been instrumental in assuring a quality product and transferring important technology for use by regulators and operators of Soviet-designed reactors. In addition, it is gratifying to observe that the diverse team learned how to surmount technical, political, and cultural barriers to effectively work together. Once Phase III is completed, work is expected to continue in Phase IV, the containment performance assessment needed for a Level 2 PRA.

Acknowledgments

The authors wish to thank their Russian colleagues and acknowledge the roles they have played in organizing, managing, and performing the work described in this paper: A. Matveev, S. Volkovitskii, and M. Mirochnitchenko, GAN; B. Gordon, T. Berg, V. Bredova, E. Zhukova, I. Kouzmina, A. Loubarski, D. Noskov, O. Safronnikova, G. Samokhine, E. Shubeiko and V. Soldatov, SEC-NRS; V. Siriapin, V. Shien, and V. Kydravtsev (deceased), EDOGP; G. Aleshin, E. Mironenko, O. Bogatov, and A. Pestrikov, KNPS; V. Kats, V. Senoedov, S. Petrunina, S. Prihodko, A. Uashkin, L. Eltsova, NIAEP; and V. Khlebitscevich, Rosenergoatom. The authors also wish to express their appreciation to T. King and T. Speis (retired) of the NRC who have provided the leadership and guidance that has made this work possible.

References

1. Azarm, M. A., et al, "Procedure Guides for a Probabilistic Risk Assessment," Draft NUREG/CR-6572, Brookhaven National Laboratory, 1998.
2. IAEA, "Procedures for Conducting Probabilistic Safety Assessments of Nuclear Power Plants (Level 2)," Safety Series No. 50-P-8, International Atomic Energy Agency, 1995.
3. NRC, "PRA Procedures Guide - A Guide to the Performance of Probabilistic Risk Assessments for Nuclear Power Plants," NUREG/CR-2300, U.S. Nuclear Regulatory Commission, September 1981.
4. Drouin, M. T., F. T. Harper, and A. L. Camp, "Analysis of Core Damage Frequency from Internal Events: Methodology, Volume 1," NUREG/CR-4550/1, Sandia National Laboratories, September 1987.

RISK COMPARISON OF PERFORMING SCHEDULED MAINTENANCE AT POWER VS. DURING SHUTDOWN

**Arthur Buslik, U.S. Nuclear Regulatory Commission
Pranab Samanta, Brookhaven National Laboratory
Bevan Staple, Sandia National Laboratories**

Results of studies of a comparison of the risk impacts of performing preventive maintenance during power operation and during periods of shutdown are presented. Two plants were studied: the PWR Surry and the BWR Grand Gulf. In the PWR, four different time windows of cold shutdown, and power operation were considered. For the BWR, the cold shutdown time windows occurring before refueling were considered, as well as power operation and the refueling plant operational states. The scope of the studies was limited to conventional internal events (excluding fire and internal floods). For the PWR, risk impacts of maintenance on the diesel generators, the auxiliary feedwater system, the low pressure injection system, and the high pressure injection system were included. For the BWR, the diesel generators, the service water system, the high pressure core spray, and the reactor core isolation cooling system were included; for the BWR case, results on taking two components out for maintenance simultaneously are included.

1.0 INTRODUCTION

This paper presents a comparison of the risk of performing scheduled maintenance at power operation and during various periods of shutdown. The motivation for the study comes from the fact that, for reasons of cost, the nuclear industry is attempting to maximize the time at power and reduce the length of refueling and maintenance outages, and consequently the nuclear industry is interested in increasing the amount of scheduled preventive maintenance performed during power operation. It is therefore of interest to determine the risk of performing preventive maintenance at power operation, and compare this risk to that of performing the maintenance during various times during shutdown and refueling. Risk-informed regulatory decision making requires this; work already performed shows that the risk during some periods of shutdown is not negligible, and hence one cannot estimate the risk change from performing maintenance at power operation instead of at shutdown by assuming that the risk from performing maintenance during shutdown is negligible. Two plants are studied, the Surry pressurized water reactor (PWR) and the Grand Gulf Nuclear Station (GGNS) boiling water reactor (BWR). The work on the PWR was performed at Brookhaven National Laboratory, with Pranab Samanta as principal investigator, and that on the BWR at Sandia National Laboratories, with Bevan Staple as principal investigator. Full reports on the PWR and BWR studies will be reported in the form of NUREG/CR reports (ref. 1 and ref. 2). Some generalizations to other plants are made, in the full studies, but will not be discussed here. The scope of the studies is limited to conventional internal events (including loss of offsite power, but excluding fires and internal floods).

Some prior work (ref. 3 and ref. 4), which stopped at estimates of the effect on core damage frequency of maintenance during shutdown had been performed, but carrying the calculations out to public risk measures yields more insights as to the relative risk of performing maintenance in different plant operating states (POSs). During shutdown the containment may be open, which affects the risk. At the same time, the time since reactor trip is greater, so that the short-lived isotopes of iodine and tellurium may have decayed appreciably, with a consequent lessening of the risk of early fatalities. The work built on prior work on the risk during low power and shutdown for Surry and Grand Gulf (ref. 5 and ref. 6).

Two types of risk measures for scheduled maintenance can be defined. The first type consists of conditional risk measures, and measures the increase in risk from taking a component, or a group of components, out for maintenance given that the plant is in a particular plant operational state (POS). One calculates the risk, given the component (or group of components) is out for maintenance and subtracts the risk, given the component (or group of components) is not in maintenance, but given that one is in the given POS. The second type of risk measure multiplies these risk measures by the duration of the scheduled maintenance performed in a calendar year on the component, or group of components, and gives the increase in the annual risk if the scheduled maintenance is done in that particular operating state. This latter type of risk measure is the most relevant, for making risk-informed decisions concerning which operational state scheduled maintenance should be performed in. Both types of risk measures were calculated in the studies.

The methodology and results of the PWR study will be discussed first, and then the results of the BWR study will be discussed.

2.0 RISK COMPARISON FOR A PWR

As noted above, the PWR considered was the Surry plant. Because it was already known that the risks were very low during refueling, when the water level is high above the reactor flange, these risks were not quantified here. Instead, the focus was on cold shutdown. The results here are affected by the fact that, at Surry, the containment was assumed open during this period. Within the cold shutdown operational state, the risk can vary with the time since shutdown, since the variation in decay heat affects system success criteria and the time available for recovery, and since the short-lived radioisotopes decay as the time from reactor trip increases. Four time windows were defined: window 1 extends from shutdown to 75 hours from shutdown; window 2 extends from 75 hours to 240 hours; window 3 extends from 240 hours to 32 days; and window 4 corresponds to times in excess of 32 days from shutdown.

In order to determine the risk impact of scheduled maintenance, a base case was defined in which all maintenance events were set to FALSE; the core damage frequency and risk measures (conditional on being in the given POS and given time window), together with their uncertainties, were then calculated. To determine the risk impact of scheduled maintenance on one or a set of components, the maintenance events corresponding to these components were set to TRUE, and the calculation repeated. The difference between the means of the core damage frequencies gives the mean conditional increase in the core damage frequency, from the component being out for maintenance in the given POS and time window. Similarly, the mean conditional increases in the public health risk measures can be calculated. However, except for the case of the diesel generators, the correlations between the uncertainty distributions for the two cases (the base case and the case with the components in maintenance) was not considered. Multiplying these mean conditional risk measures by the assumed time per year for

maintenance in the given POS gives the increase in the calendar year risk measures. For more details concerning the methodology used, see ref. 1; ref. 7 gives details on the cold shutdown model used. In this paper, results will be given only for preventive maintenance on single components, not simultaneous maintenance on two components. Some results on the simultaneous maintenance of two components is given in ref. 1.

The base case results for the core damage frequency (CDF), total early fatalities, and total latent cancer fatalities, are given in Table 1 below. These values are conditional on being in the given plant operational state and time window; the units are per year. These base case results, as already noted, were based on zero maintenance unavailability for all components; although one can use this base case to see the risk impact of maintenance on a given component or set of components, the value of the core damage frequency with average maintenance unavailabilities would be higher.

The core damage frequency (conditional on being in the given POS and time window) is greater in time windows 2 and 3 than in time window 1, primarily because the probability of containment sump plugging was assumed to be greater in time windows 2 and 3. This is a consequence of the greater amount of maintenance inside containment performed in these time windows; the material used could be swept to the containment sump and plug it, when emergency cooling recirculation was required.

Table 1. Mean base case results for Surry

| | At Power | Cold Shutdown | | | |
|-------------------------------|----------|---------------|----------|----------|----------|
| | | Window 1 | Window 2 | Window 3 | Window 4 |
| CDF (1/yr) | 4.3E-5 | 2.0E-5 | 4.4E-5 | 4.1E-5 | 1.2E-5 |
| Tot. Early Fatalities (1/yr) | 7.0E-7 | 1.3E-7 | 3.1E-8 | 9.9E-10 | 3.7E-11 |
| Tot. Cancer Fatalities (1/yr) | 1.1E-2 | 3.2E-2 | 3.8E-2 | 2.6E-2 | 8.5E-3 |

The components, or component trains, for which results are presented in this paper are:

- TD-AFW Turbine-driven pump in the Auxiliary Feedwater System (AFW)
- MD-AFW Motor-driven pump in the AFW
- DG1 Diesel Generator 1, dedicated to unit 1
- DG3 Diesel Generator 3, swing diesel, shared between both units
- LPI-B Low pressure injection pump, train B
- HPI-B High pressure injection pump, train B
- CW-B Cooling water train B
- SW-B Service water train B

Table 2 gives the increase in the calendar year mean CDF from preventive maintenance (PM) of 7 days per year, on each of the components listed, at power, and during the four time windows of cold shutdown. Table 3 gives the increase in the calendar year total early fatalities from PM of 7 days per year, and Table 4 gives the corresponding results for total latent cancer fatalities.

Table 2. Increase in the calendar year mean CDF from PM of 7 days/year, at Surry

| | Power Operation | Window 1 | Window 2 | Window 3 | Window 4 |
|--------|--------------------|----------|----------|----------|----------|
| DG1 | 6.3E-06 | 8.2E-07 | 5.7E-07 | 4.3E-07 | 3.8E-07 |
| DG3 | 3.5E-06 | 4.1E-07 | 2.7E-07 | 2.0E-07 | 1.7E-07 |
| TD-AFW | 1.6E-06 | — | — | — | — |
| LPI-B | 3.8 E-07 | 1.8E-07 | 1.2E-07 | 1.2E-07 | 1.2E-07 |
| SW-B | 3.0E-07 | 4.2E-08 | 3.5E-08 | 1.9E-09 | 3.8E-09 |
| MD-AFW | 2.0E-07 | 2.1E-08 | 4.8E-08 | 4.4E-08 | 1.2E-08 |
| CW-B | 1.4E-07 | 3.1E-08 | 2.3E08 | 1.9E-09 | 3.8E-08 |
| HPI-B | 2.5E-08 | 2.3E-08 | 1.7E-08 | 1.0E-09 | 9.6E-10 |

Table 3 Increase in yearly total early fatalities from PM of 7 days/year, at Surry

| | Power Operation | Window 1 | Window 2 | Window 3 | Window 4 |
|--------|--------------------|----------|----------|----------|----------|
| DG1 | 2.5E-09 | 1.9E-09 | 1.2E-10 | 5.4E-12 | 2.2E-13 |
| DG3 | 1.4E-09 | 9.2E-10 | 5.0E-11 | 2.5E-12 | 1.0E-13 |
| TD-AFW | 6.9E-10 | — | — | — | — |
| LPI-B | ε | 1.0E-09 | 8.6E-11 | 2.3E-12 | 1.3E-13 |
| SW-B | 2.1E-10 | 2.3E-10 | 2.1E-11 | 3.8E-14 | ε |
| MD-AFW | ε | 5.8E-11 | 2.9E-11 | 7.7E-13 | ε |
| CW-B | 1.7E-10 | 1.5E-10 | 1.5E-11 | 5.8E-14 | ε |
| HPI-B | 3.8E-11 | 1.2E-10 | 1.2E-11 | 1.9E-14 | ε |

ε: negligible contribution

Table 4 Increase in yearly total latent cancer fatalities from PM of 7 days/year, at Surry

| | Power Operation | Window 1 | Window 2 | Window 3 | Window 4 |
|--------|--------------------|----------|----------|----------|----------|
| DG1 | 2.2E-04 | 6.7E-04 | 3.2E-04 | 1.9E-04 | 6.7E-05 |
| DG3 | 1.2E-04 | 3.4E-04 | 1.5E-04 | 9.4E-05 | 3.0E-05 |
| TD-AFW | 4.8E-05 | — | — | — | — |
| LPI-B | 5.8E-06 | 1.7E-04 | 7.9E-05 | 5.4E-05 | 2.2E-05 |
| SW-B | 1.1E-04 | 3.8E-05 | 1.9E-05 | 1.9E-06 | 1.9E-07 |
| MD-AFW | 7.7E-06 | 3.8E-06 | 2.3E-05 | 1.7E-05 | € |
| CW-B | 1.0E-04 | 2.88E-05 | 1.5E-05 | 1.9E-06 | 3.8E-07 |
| HPI-B | 2.7E-05 | 2.3E-05 | 1.2E-05 | 1.9E-06 | € |

€: negligible contribution

The Surry plant, as modeled in the study, consists of two units, each with a dedicated diesel, and a third diesel shared between the two units. Maintenance on the Unit 1 dedicated diesel generator (DG1) has a greater risk impact (on Unit 1) than maintenance on the shared diesel generator (DG3). For DG1, the increase in the yearly core damage frequency is 6E-6 per year, for seven days per year of maintenance at power, while in window 4 it is 4E-7 per year, more than an order of magnitude lower than at power. The increase in expected yearly total latent fatalities from 7 days per year maintenance on DG1 at power is 2.2E-4 per year, while during time window 4 the corresponding increase is 6.7E-5 per year. For DG3, the risk impact of maintenance on Unit 1 is less, but one also has to consider the risk impact on Unit 2. Also, since it is unlikely that both units will be shut down for maintenance or refueling at the same time, the most likely risk reduction strategy would be to perform maintenance on DG3 when one unit is in time window 4 of cold shutdown or in the refueling POS, and the other unit is at power. Maintenance on diesel generators at power has a greater risk impact than maintenance on the other components studied.

The turbine-driven auxiliary feedwater pump (TD-AFW) is not used during shutdown; since seven days of maintenance per year on the TD-AFW at power results in an increase in the yearly core damage frequency of 1.6E-6 per year, and an increase in expected total latent cancer fatalities of 5E-5 per year, it may be desirable to perform maintenance on the TD-AFW during cold shutdown.

For a low pressure injection pump train, the increase in the yearly core damage frequency from 7 days per year of maintenance at power is 3.8E-7 per year, while the corresponding value for time window 4 of cold shutdown is 1.2E-7 per year. The corresponding increase in total latent cancer fatalities is 5.8E-6 per year for maintenance at power, and 2.2E-5 per year for maintenance in time window 4 of cold shutdown. Because of the increased health effects from maintenance performed in cold shutdown on a low pressure injection pump train, and because only a limited amount of maintenance can be performed while in the refueling plant operational state, it would appear desirable to perform maintenance on the low pressure

injection pump at power, especially since there is probably a cost saving in so doing.

For each of the time-windows of cold shutdown, the relative ranking of components that cause the largest risk impacts are similar: diesel generator, low pressure injection train, service water pump train, motor-driven auxiliary feedwater pump train, component cooling water pump train, and high pressure injection pump train. Comparison of the impact across time-windows shows that early fatalities are reduced significantly for time-windows 3 and 4, i.e., when the elapsed time from shutdown is longer, because of the decay of short-lived isotopes. Reduction in latent cancer fatalities is less pronounced as the reduction here depends primarily on the reduction in the core damage frequency.

3.0 RISK COMPARISON FOR A BWR

As already noted, the BWR considered was Grand Gulf. The study considered the following POSs: full-power operation (POS 0), cold shutdown (POS 5), refueling with vessel water level at the steam lines (POS 6), and refueling with the vessel flooded up to the upper containment pool and the refueling transfer tube open (POS 7). For POS 5, only the time windows before the refueling POS were considered. The components considered were the emergency diesel generators (EDGs), the standby service water (SSW) system motor-driven pumps (MDPs), the reactor core isolation cooling (RCIC) system turbine-driven pump (TDP), and the high pressure core spray (HPCS) system MDP. Maintenance on single components and selected pairs of components was considered.

3.1 Methodology

The conditional risk measures considered were:

- (1) *Increase in Conditional Core Damage Frequency, I_{CDF} .* $I_{CDF}=C_1-C_0$, where C_1 is the core damage frequency, conditional on being in a given POS, and conditional on the component(s) being out for maintenance; C_0 is the core damage frequency, conditional on being in the same POS, but with the component(s) not being out for maintenance.
- (2) *Increase in Conditional Individual Early Fatality Risk (IEFR), I_{IEFR} .* $I_{IEFR}=E_1-E_0$, where E_1 is the IEFR conditional on the plant being in a given POS and conditional on the component(s) being out for maintenance and E_0 is the IEFR in the same POS but with the component(s) not being out for maintenance. The individual early fatality risk is the probability, per unit time in the given POS, of an individual within one mile of the plant's exclusion boundary dying within one year of an accident, from early exposure to radionuclides released following an accident.
- (3) *Increase in Conditional Individual Latent Cancer Fatality Risk (ILCFR), I_{ILCFR} .* $I_{ILCFR}=F_1-F_0$, where F_1 is the ILCFR conditional on the plant being in a given POS and conditional on the component(s) being out for maintenance and F_0 is the ILCFR in the same POS but with the component(s) not being out for maintenance. The individual latent cancer fatality risk is the probability, per unit time of operation in the POS, of dying from cancer for an individual within 10 miles of the plant due to both early and chronic (i.e., more than 7 days after the accident) exposure to radionuclides released following an accident.
- (4) *Increase in Conditional Population Dose Risk within 50 miles (PDR50), I_{PDR50} .* $I_{PDR50}=D_1-D_0$, where

D_1 is the PDR50 conditional on the plant being in a given POS and conditional on the component(s) being out for maintenance and D_0 is the PDR50 in the same POS but with the component(s) not being out for maintenance. The population dose, expressed in effective dose equivalents for whole body exposure, occurs due to both early and chronic exposure pathways to the population within 50 miles of the reactor, from the radionuclides released following an accident.

Risk measures representing the yearly increase in risk due to the performance of preventive maintenance on a component or group of components were also calculated. These risk measures are the product of the measures given above and the time spent per year in maintenance, while in a given POS. This group of risk measures consists of:

- (1) *Yearly Core Damage Frequency Contribution, I_{YCDF}* . The I_{YCDF} is the expected increase in the calendar year CDF due to the performance of preventive maintenance on the component(s) (i.e., the product of the I_{CDF} and the maintenance duration per year, measured as the fraction of the year spent performing preventive maintenance on the component(s)).
- (2) *Yearly Individual Early Fatality Risk Contribution, I_{YIEFR}* . The I_{YIEFR} is the expected increase in the yearly number of early fatalities due to the performance of preventive maintenance on the component(s) (i.e., the product of the I_{IEFR} and the fraction of the year spent performing preventive maintenance on the component).
- (3) *Yearly Individual Latent Cancer Fatality Risk Contribution, I_{YILCFR}* . The I_{YILCFR} is the expected increase in the yearly number of latent cancer fatalities per calendar year due to the performance of preventive maintenance on the component(s) (i.e., the product of the I_{ILCFR} and the fraction of the year spent performing preventive maintenance on the component).
- (4) *Yearly Population Dose Risk Contribution, I_{YPDR50}* . The I_{YPDR50} is the expected increase in the yearly population dose within 50 miles due to the performance of preventive maintenance on the component(s) (i.e., the product of the I_{PDR50} and the fraction of the year spent performing preventive maintenance on the component).

The concept of degree of belief, D , was also introduced as a means of supporting the decision on when to perform maintenance when there are overlaps in the uncertainty distributions of the results across different POSs. The degree of belief represents the percentage of the uncertainty observations (i.e., sample points) for which a risk measure is greater during a selected shutdown POS than at power. To add qualitative worth to the degree of belief concept, ranges were established for the percentages and expressed subjectively as follows:

- $D < 20\%$ → very low degree of belief
- $20\% < D < 40\%$ → low degree of belief
- $40\% < D < 60\%$ → neutral degree of belief
- $60\% < D < 80\%$ → medium degree of belief
- $80\% < D < 100\%$ → high degree of belief

To illustrate the concept, consider the case where POS 0 and POS 5 have comparable I_{CDF} mean values but in only 15% of the sample points is the I_{CDF} greater during POS 5 than during POS 0. Even with

comparable means, this very low degree of belief (i.e., 15%) that the I_{CDF} is higher during POS 5 than during POS 0 provides additional impetus to perform preventive maintenance during POS 5 rather than during POS 0. On the other hand a D of 55% would indicate a more neutral degree belief thus allowing decision makers to choose either POS or perhaps base their selection on other factors such as economics.

Level 1, 2, and 3 PRA models were developed for GGNS to simulate accident conditions both at power and at shutdown. The Level 1 models for POSs 0 and 5 are modifications of existing models for GGNS generated in previous NRC-sponsored studies (ref. 6 for POS 5 and ref. 8 for POS 0). The Level 1 models for refueling (POSs 6 and 7) were developed specifically for this study. The Level 2/3 models are adaptations of existing models for power operation (ref. 9) modified to include features specific to low-power and shutdown operations.

The Level 1 PRA models for each POS were integrated into a single global model with a common database. The integration assured that the same sets of sample points were used in the uncertainty analysis used in evaluating the CDF for each POS. A separate uncertainty analysis was performed on the Level 2/3 models. Thus the Level 2/3 models for each POS were also integrated into a single global model. This step was required to obtain the uncertainty distributions for the CDF and public risk measures calculated for each POS. This method retains the correlations between the base case and the case where components are out for maintenance, when the risk increases from maintenance are calculated.

The Level 1 models for GGNS were linked to the Level 2/3 models through the use of plant damage states (PDSs), using a computer program designed for the purpose. This linking allowed sensitivities in the Level 1 models to be quickly observed in the Level 2/3 results.

The risk-based measures were evaluated for several components at GGNS using the linked/integrated PRA models. The methodology was demonstrated for preventive maintenance activities on the emergency diesel generators (EDGs), the standby service water (SSW) system motor-driven pumps (MDPs), the reactor core isolation cooling (RCIC) system turbine-driven pump (TDP), and the high pressure core spray (HPCS) system MDP. Preventive maintenance on single components and selected combinations of two components were performed. The results of these evaluations led to insights about the risk-impact of performing preventive maintenance at power versus shutdown.

The final step in the methodology was to extend the insights from the GGNS evaluation to additional boiling water reactors (BWRs). Since there are few published shutdown studies for BWRs, this comparison was limited to the risk impact of preventive maintenance on single components during power operations. These results will not be presented in this paper; see ref. 2.

3.2 Results

Baseline CDF and Public Risk Results

The PRA models for GGNS were evaluated to obtain baseline CDF and public risk values. As indicated in Table 5, the mean baseline CDF, IEFr, ILCFR, and PDR50 during cold shutdown are of the same order of magnitude as the corresponding values during full power operations. However, the values during refueling are one to two orders of magnitude less than those during full power operations. These

results can be attributed to the fact that during the early stages of cold shutdown the decay heat is still significant compared to power and there are similarities in response to the dominant of accidents types (e.g., loss of offsite power). However, during refueling the decay heat is significantly reduced compared to full power operation so the time available to respond to certain core damage accidents is substantially lengthened and the energy removal requirements are reduced. In addition, there is a large inventory of water available in the upper containment pools for accident mitigation (during POS 7) and even though the vessel head and the containment is open, the amounts of short-lived radionuclides contributing to early fatality risk are much smaller than those during full power operation (i.e., the radionuclides have decayed).

Table 5. Mean baseline risk measures at Grand Gulf.^a

| POS | Mean CDF (yr ⁻¹) | Mean IEFR (IEF/yr) | Mean ILCFR (ILCF/yr) | Mean PDR50 (Sv/yr) |
|-----|---------------------------------|-----------------------|-------------------------|-----------------------|
| 0 | 9.8E-6 | 1.5E-10 | 2.8E-9 | 2.0E-2 |
| 5 | 1.4E-5 | 5.8E-10 | 6.0E-9 | 5.8E-2 |
| 6 | 6.7E-7 | 1.1E-11 | 3.7E-10 | 3.4E-3 |
| 7 | 3.8E-7 | 8.9E-12 | 2.1E-10 | 2.0E-3 |

^a Values are shown conditional on being in the given POS (i.e., does not include fraction of year in the POS).

Conditional Risk Increase From the Scheduling of Preventive Maintenance

As indicated in Table 6, the mean I_{CDF} , I_{IEFR} , I_{ILCFR} , and I_{PDR50} from performing preventive maintenance activities on single components (the EDGs, the SSW MDPs, or the HPCS MDP) during cold shutdown are of the same order of magnitude as the mean I_{CDF} , I_{IEFR} , I_{ILCFR} , and I_{PDR50} from performing similar activities during power operations. In addition, there are significant overlaps in the distributions of the measures in both POSs. This is characterized by neutral to medium degrees of belief that observation of either the I_{CDF} , the I_{IEFR} , the I_{ILCFR} , or the I_{PDR50} will be higher during cold shutdown than during full power operations. The motor driven pump B of the SSW system can be used to inject water into the vessel through a crosstie to train B of the low pressure injection system; this accounts for some of the asymmetry between the SSW pumps in the table.

The mean risk measures from performing preventive maintenance activities on the single components during refueling are one to two orders of magnitude less than those from performing similar maintenance activities during power operations. In addition, there is usually little or no overlap in the distributions of the measures in both POSs and very low degrees of belief that any observation of either the I_{CDF} , the I_{IEFR} , the I_{ILCFR} or the I_{PDR50} will be higher during refueling than during power operations. The lower risk impacts from performing preventive maintenance during refueling can be attributed to the following factors: (1) the decay heat is significantly reduced compared to power operation so the time available for response to certain core damage accidents is substantially lengthened and the energy removal requirements are reduced during refueling, (2) there is a large inventory of water available in the upper

containment pools for accident mitigation especially during POS 7, and (3) even though the vessel head is removed and the containment is open, the amounts of short-lived radionuclides contributing to early fatality risk are substantially smaller than those during power operation due to radionuclide decay. Maintenance on diesel generators in POS 7 has zero impact because the amount of water above the reactor core is so great that the time to boiloff in station blackout sequences exceeds the mission time of 24 hours used in the study.

The risk measures associated with performing multiple system intra-division maintenance activities (i.e., maintenance on multiple components within a single division while avoiding maintenance on redundant divisions) are shown in Table 7. A comparison of the risk measures for the different POSs indicates the same trend as identified for the single components. The risk measures for the multiple component maintenance outages range from approximately a factor of two to ten times the values obtained from performing the single component maintenance activities. The risk measures for the multiple component maintenance outages also show levels of overlap in the risk distributions and degrees of belief similar to those for the single component maintenance activities.

From these observations it can be concluded that considering uncertainties in the results, there appears to be no significant risk advantage from performing preventive maintenance activities on the EDGs, the SSW MDPs, or the HPCS MDP at GGNS (either singularly or in the combinations evaluated) during cold shutdown instead of at power. Thus, the decision on whether to perform preventive maintenance on these components during power operation or during cold shutdown should probably be based on other factors such as economics. On the other hand, there does appear to be a significant risk advantage from performing preventive maintenance activities on these components (both singularly or in the combinations evaluated) during refueling versus during power operations. Thus, preventive maintenance of long duration on these components could be scheduled for refueling while those of shorter duration could be scheduled during power operation or cold shutdown.

Yearly Risk Contributions From the Scheduling of Preventive Maintenance

The yearly risk contributions from performing preventive maintenance on selected single components at GGNS are provided in Table 8. These values were calculated assuming that the preventive maintenance outages would be of the same duration for each POS. Generally, the mean I_{YCDF} , I_{YIEFR} , I_{YILCFR} , and I_{YPDR50} from performing preventive maintenance activities on single components of the EDGs, the SSW MDPs, the HPCS MDPs, or the RCIC TDP during power operations, cold shutdown, and refueling constitute less than a 4% increase in the corresponding mean baseline risks (i.e., the mean CDF, IEFR, ILCFR, and PDR50 presented in Table 1) during these plant states.

The yearly risk contributions from performing preventive maintenance on selected combinations of components at GGNS are provided in Table 9 for each POS. These values also were calculated assuming that the preventive maintenance outages would be of the same duration for each POS. The results of performing multiple system intra-division maintenance activities on the EDGs, the SSW MDPs, and the HPCS MDPs are similar to the results for the single components. This fact can be attributed to two factors. First, GGNS's maintenance practices avoids taking redundant divisions of equipment out for maintenance and thus moderates the risk increases associated with preventive maintenance. Second, many maintenance activities involve maintaining front-line systems and their respective support systems during the same time period. The unavailability of the support systems makes the front-line systems

unavailable.

The results in Tables 8 and 9 indicates that maintaining both front-line systems and support systems at the same time does not always create significant increases in plant risks. For example, SSW MDP C provides cooling to EDG 3, therefore, failure of SSW MDP C fails EDG 3. Thus, taking SSW MDP C and EDG 3 out at the same time for preventive maintenance does not create significant increases in risks above those created by taking only SSW MDP C out for maintenance. Note, however, that the yearly risk contributions associated with preventive maintenance on both SSW MDP C and EDG 3 is less than the values when only EDG 3 is taken out for maintenance. This occurs because the assumed downtime for preventive maintenance for just EDG 3 is $1.2E-2$ per year, while the assumed downtime associated with performing preventive maintenance on both of the components at the same time is $1.7E-3$, the minimum of the downtimes of the two components. Different results would be obtained for different maintenance down times.

From these observations it can be concluded that under the current preventive maintenance practices at GGNS, single train and multiple intra-division preventive maintenance activities on the EDGs, the SSW MDPs, the HPCS MDP, and the RCIC TDP during any POS do not induce significant risks above the mean baseline risk. However, a significant increase in the downtime or the frequency of preventive maintenance activities on the components beyond current practice could appreciably increase the plant's mean baseline risks. Therefore, managing the scheduling, the duration, and the frequency of preventive maintenance outages should be considered an integral part of the plant's maintenance plan.

4.0 REFERENCES

1. J. W. Yang, T-L. Chu, G. Martinez-Guridi, and P.K. Samanta, "Risk Comparison of Scheduling Preventive Maintenance During Shutdown vs. During Power Operation for PWRs", NUREG/CR-6616, to be published.
2. B.D. Staple, T. Brown, J. Gregory, "Risk Comparisons of Scheduling Preventive Maintenance for Boiling Water Reactors During Shutdown and Power Operations", to be published as an NUREG/CR report.
3. P.K. Samanta et al., "Emergency Diesel Generator Maintenance and Failure Unavailability, and their Risk Insights," NUREG/CR-5994, October 1994.
4. B.D. Staple et al., "Risk Impact of BWR Technical Specifications Requirements During Shutdown", NUREG/CR-6166, September 1994.
5. T.L. Chu and W.T. Pratt (Editors), "Evaluation of Potential Severe Accidents During Low Power and Shutdown Operations at Surry, Unit 1", NUREG/CR-6144, October 1995.
6. D.W. Whitehead et al., "Evaluation of Potential Severe Accidents During Low Power and Shutdown Accident Operations at Grand Gulf, Unit 1", NUREG/CR-6143, July 1994.
7. T.L. Chu, J.W. Yang, G. Martinez-Guridi, and P.K. Samanta, "A Simplified Level-1 Probabilistic Risk Assessment Model of a PWR in a Cold Shutdown Condition", to be published as an

NUREG/CR report.

8. M.T. Drouin, et al., "Analysis of Core Damage Frequency: Grand Gulf Unit 1 Internal Events," NUREG/CR-4550, Vol. 6, Rev. 1, September 1989.
9. T.D. Brown et al., "Evaluation of Severe Accident Risks: Grand Gulf Unit 1", NUREG/CR-4551, vol. 6, rev. 1, December 1990.

Table 6. Conditional risk increase from preventive maintenance on single components.*

| Mean I_{CDF} (yr^{-1}) | | | | | | | |
|------------------------------|--------------|--------------|--------------|--------------|--------------|--------------|----------|
| POS | EDG 1 | EDG 3 | SSW MDP A | SSW MDP B | SSW MDP C | HPCS MDP | RCIC TDP |
| POS 0 | 2.8E-5 | 3.2E-5 | 2.5E-5 | 2.7E-5 | 2.8E-5 | 2.0E-6 | 1.1E-4 |
| POS 5 | 1.8E-5 (42) | 2.1E-5 (42) | 2.1E-5 (54) | 2.0E-5 (39) | 2.4E-5 (60) | 2.0E-6 (33) | 0 |
| POS 6 | 1.6E-7 (0) | 1.6E-7 (0) | 0 | 3.8E-6 (22) | 1.7E-7 (1) | 1.5E-7 (8) | 0 |
| POS 7 | 0 | 0 | 0 | 2.0E-6 (15) | 0 | 0 | 0 |
| Mean I_{IEFR} (IEF/yr) | | | | | | | |
| POS 0 | 3.8E-10 | 4.5E-10 | 3.6E-10 | 3.9E-10 | 4.1E-10 | 1.7E-11 | 1.6E-9 |
| POS 5 | 8.2E-10 (50) | 9.6E-10 (52) | 9.9E-10 (57) | 9.9E-10 (52) | 1.3E-9 (70) | 1.9E-10 (83) | 0.0 |
| POS 6 | 9.6E-13 (1) | 9.6E-13 (1) | 0.0 | 7.9E-11 (19) | 8.9E-14 (0) | 7.1E-14 (2) | 0.0 |
| POS 7 | 0.0 | 0.0 | 0.0 | 3.8E-11 (16) | 0.0 | 0.0 | 0.0 |
| Mean I_{ILCFR} (ILCF/yr) | | | | | | | |
| POS 0 | 7.4E-9 | 8.6E-9 | 6.5E-9 | 7.2E-9 | 7.4E-9 | 6.2E-10 | 3.0E-8 |
| POS 5 | 9.0E-9 (57) | 1.0E-8 (56) | 1.0E-8 (67) | 1.0E-8 (53) | 1.2E-8 (71) | 7.8E-10 (45) | 0.0 |
| POS 6 | 7.6E-11 (0) | 7.6E-11 (0) | 0.0 | 2.2E-9 (34) | 6.4E-11 (30) | 5.5E-11 (9) | 0.0 |
| POS 7 | 0.0 | 0.0 | 0.0 | 1.1E-9 (26) | 0.0 | 0.0 | 0.0 |
| Mean I_{PDRS} (Sv/yr) | | | | | | | |
| POS 0 | 5.1E-2 | 6.0E-2 | 4.4E-2 | 5.0E-2 | 5.1E-2 | 4.3E-3 | 2.0E-1 |
| POS 5 | 9.2E-2 (61) | 9.4E-2 (62) | 9.3E-2 (70) | 9.2E-2 (56) | 1.2E-1 (78) | 1.4E-2 (60) | 0.0 |
| POS 6 | 6.1E-4 (0) | 6.1E-4 (0) | 0.0 | 1.9E-2 (36) | 3.3E-4 (1) | 2.8E-4 (8) | 0.0 |
| POS 7 | 0.0 | 0.0 | 0.0 | 9.6E-3 (29) | 0.0 | 0.0 | 0.0 |

* Values in the parentheses are percentages and represent the degree of belief that observations of that measure are higher during that POS than during full power operation. Risk-impact values of 0.0 implies that the PRA model for that POS has no cut sets with maintenance on that equipment.

Table 7. Conditional risk increase from preventive maintenance on multiple components.*

| Mean I_{CDF} (yr^{-1}) | | | |
|------------------------------|---------------------|---------------------|--------------------|
| POS | EDG 3 and SSW MDP C | EDG 2 and SSW MDP B | EDG 3 and HPCS MDP |
| POS 0 | 7.2E-5 | 6.5E-5 | 6.4E-5 |
| POS 5 | 4.2E-5 (49) | 3.6E-5 (38) | 2.1E-5 (39) |
| POS 6 | 2.9E-7 (1) | 3.7E-6 (16) | 2.7E-7 (0) |
| POS 7 | 0.0 | 2.0E-6 (9) | 0.0 |
| Mean I_{IEFR} (IEF/ yr) | | | |
| POS 0 | 1.3E-9 | 1.1E-9 | 1.1E-9 |
| POS 5 | 2.1E-9 (64) | 1.6E-9 (53) | 1.1E-9 (63) |
| POS 6 | 1.4E-12 (1) | 7.6E-11 (14) | 1.4E-12 (1) |
| POS 7 | 0.0 | 4.9E-11(10) | 0.0 |
| Mean I_{ILCFR} (ILCF /yr) | | | |
| POS 0 | 2.0E-8 | 1.8E-8 | 1.6E-8 |
| POS 5 | 2.0E-8 (66) | 1.7E-8 (57) | 9.9E-9 (58) |
| POS 6 | 1.2E-10 (1) | 2.1E-9 (24) | 1.1E-10 (1) |
| POS 7 | 0.0 | 1.1E-9 (16) | 0.0 |
| Mean I_{PDR50} Sv /yr) | | | |
| POS 0 | 1.5E-1 | 1.3E-1 | 1.2E-1 |
| POS 5 | 1.9E-1 (73) | 1.6E-1 (61) | 9.9E-2 (60) |
| POS 6 | 8.0E-4 (1) | 1.9E-2 (25) | 7.5E-4 (81) |
| POS 7 | 0.0 | 1.0E-2 (19) | 0.0 |

* Values in the parentheses are percentages and represent the degree of belief that observations of that measure are higher during that POS than during full power operation. Risk-impact values of 0.0 implies that the PRA model for that POS has no cut sets with maintenance on that equipment.

Table 8. Yearly risk contribution from preventive maintenance on single components.^a

| Mean I_{VCOF} (yr^{-1}) | | | | | | | |
|-------------------------------|---------|---------|-----------|-----------|-----------|----------|----------|
| POS | EDG 1 | EDG 3 | SSW MDP A | SSW MDP B | SSW MDP C | HPCS MDP | RCIC TDP |
| POS 0 | 3.4E-7 | 3.8E-7 | 4.3E-8 | 4.6E-8 | 4.8E-8 | 1.2E-8 | 5.1E-7 |
| POS 5 | 2.2E-7 | 2.5E-7 | 3.6E-8 | 3.4E-8 | 4.1E-8 | 1.2E-8 | 0.0 |
| POS 6 | 1.9E-9 | 1.9E-9 | 0.0 | 6.5E-9 | 2.9E-10 | 9.2E-10 | 0.0 |
| POS 7 | 0.0 | 0.0 | 0.0 | 3.4E-9 | 0.0 | 0.0 | 0.0 |
| Mean I_{VIEFR} (IEF/yr) | | | | | | | |
| POS 0 | 4.6E-12 | 5.4E-12 | 6.1E-13 | 6.6E-13 | 7.0E-13 | 1.0E-13 | 7.4E-12 |
| POS 5 | 9.8E-12 | 1.2E-11 | 1.7E-12 | 1.7E-12 | 2.2E-12 | 1.2E-12 | 0.0 |
| POS 6 | 1.2E-14 | 1.2E-14 | 0.0 | 1.3E-13 | 1.5E-16 | 4.3E-16 | 0.0 |
| POS 7 | 0.0 | 0.0 | 0.0 | 6.5E-14 | 0.0 | 0.0 | 0.0 |
| Mean I_{VILCFR} (ILCF/yr) | | | | | | | |
| POS 0 | 8.9E-11 | 1.0E-10 | 1.1E-11 | 1.2E-11 | 1.3E-11 | 3.8E-12 | 1.4E-10 |
| POS 5 | 1.1E-10 | 1.2E-10 | 1.7E-11 | 1.7E-11 | 2.0E-11 | 4.8E-12 | 0.0 |
| POS 6 | 9.1E-13 | 9.1E-13 | 0.0 | 3.7E-12 | 1.1E-13 | 3.4E-13 | 0.0 |
| POS 7 | 0.0 | 0.0 | 0.0 | 1.9E-12 | 0.0 | 0.0 | 0.0 |
| Mean I_{VPDR50} (Sv/yr) | | | | | | | |
| POS 0 | 6.1E-4 | 7.2E-4 | 7.5E-5 | 8.5E-5 | 8.7E-5 | 2.6E-5 | 9.2E-4 |
| POS 5 | 1.1E-3 | 1.1E-3 | 1.6E-4 | 1.6E-4 | 2.0E-4 | 8.5E-5 | 0.0 |
| POS 6 | 7.3E-6 | 7.3E-6 | 0.0 | 3.2E-5 | 5.6E-7 | 1.7E-6 | 0.0 |
| POS 7 | 0.0 | 0.0 | 0.0 | 1.6E-5 | 0.0 | 0.0 | 0.0 |

^a Yearly risk-impact measure are calculated based on preventive maintenance durations (measured in fractions of a year) of 1.2E-2, 6.1E-3, 4.6E-3, and 1.7E-3 for the EDGs, HPCS MDP, RCIC TDP, and SSW MDPs, respectively. Risk-impact values of 0.0 implies that the PRA model for that POS has no cut sets with maintenance on that equipment.

Table 9. Yearly risk contribution from preventive maintenance on multiple components.^a

| Mean I_{YCDF} (yr^{-1}) | | | |
|--|----------------------------|----------------------------|---------------------------|
| POS | EDG 3 and SSW MDP C | EDG 2 and SSW MDP B | EDG 3 and HPCS MDP |
| POS 0 | 1.2E-7 | 1.1E-7 | 3.9E-7 |
| POS 5 | 7.1E-8 | 6.1E-8 | 1.3E-7 |
| POS 6 | 4.9E-10 | 6.3E-9 | 1.7E-9 |
| POS 7 | 0.0 | 3.4E-9 | 0.0 |
| Mean I_{YIEFR} (IEF/ yr) | | | |
| POS 0 | 2.2E-12 | 1.9E-12 | 6.7E-12 |
| POS 5 | 3.6E-12 | 2.7E-12 | 6.7E-12 |
| POS 6 | 2.4E-15 | 1.3E-13 | 8.5E-15 |
| POS 7 | 0.0 | 8.3E-14 | 0.0 |
| Mean I_{YILCFR} (ILCF /yr) | | | |
| POS 0 | 3.4E-11 | 3.1E-11 | 9.8E-11 |
| POS 5 | 3.4E-11 | 2.9E-11 | 6.0E-11 |
| POS 6 | 2.0E-13 | 3.6E-12 | 6.7E-13 |
| POS 7 | 0.0 | 1.9E-12 | 0.0 |
| Mean I_{YVDR50} (Sv /yr) | | | |
| POS 0 | 2.6E-4 | 2.2E-4 | 7.3E-4 |
| POS 5 | 3.2E-4 | 2.7E-4 | 6.0E-4 |
| POS 6 | 1.4E-6 | 3.2E-5 | 4.6E-6 |
| POS 7 | 0.0 | 1.7E-5 | 0.0 |
| ^a Yearly risk-impact measure for multiple component outages are calculated based on the component with the shortest preventive maintenance unavailability since that would be the largest possible outage time for both components. A value of 1.7E-3 was used for outage combinations involving EDG 3 and SSW MDP C and EDG 2 and SSW MDP B, and a value of 6.1E-3 was used for the combination of EDG 3 and HPCS MDP out for maintenance. | | | |

BIBLIOGRAPHIC DATA SHEET

(See instructions on the reverse)

2. TITLE AND SUBTITLE

Proceedings of the Twenty-Sixth Water Reactor Safety Information Meeting
Digital Instrumentation and Control, The Halden Program, Structural Performance, and PRA
Methods and Applications

5. AUTHOR(S)

Compiled by Susan Monteleone, Brookhaven National Laboratory

6. PERFORMING ORGANIZATION - NAME AND ADDRESS (If NRC, provide Division, Office or Region, U.S. Nuclear Regulatory Commission, and mailing address; if contractor, provide name and mailing address.)

Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555-0001

9. SPONSORING ORGANIZATION - NAME AND ADDRESS (If NRC, type "Same as above"; if contractor, provide NRC Division, Office or Region, U.S. Nuclear Regulatory Commission, and mailing address.)

Same as 8. above.

10. SUPPLEMENTARY NOTES

S. Nesmith, NRC Project Manager Proceedings prepared by Brookhaven National Laboratory

11. ABSTRACT (200 words or less)

This three-volume report contains papers presented at the Twenty-Sixth Water Reactor Safety Information Meeting held at the Bethesda Marriott Hotel, Bethesda, Maryland, October 26-28, 1998. The papers are printed in the order of their presentation in each session and describe progress and results of programs in nuclear safety research conducted in this country and abroad. Foreign participation in the meeting included papers presented by researchers from France, Germany, Italy, Japan, Norway, Russia, Sweden and Switzerland. The titles of the papers and the names of the authors have been updated and may differ from those that appeared in the final program of the meeting.

12. KEY WORDS/DESCRIPTORS (List words or phrases that will assist researchers in locating the report.)

BWR Type Reactors - Reactor Safety, Nuclear Power Plants - Reactor Safety, PWR Type Reactors -
Reactor Safety, Reactor Safety - Meetings

1. REPORT NUMBER
(Assigned by NRC, Add Vol., Supp., Rev.,
and Addendum Numbers, if any.)

NUREG/CP-0166
Vol. 2

3. DATE REPORT PUBLISHED

| MONTH | YEAR |
|-------|------|
| June | 1999 |

4. FIN OR GRANT NUMBER

A3988

6. TYPE OF REPORT

Conference Proceedings

7. PERIOD COVERED (Inclusive Dates)

October 26-28, 1998

13. AVAILABILITY STATEMENT

unlimited

14. SECURITY CLASSIFICATION

(This Page)

unclassified

(This Report)

unclassified

15. NUMBER OF PAGES

16. PRICE



Federal Recycling Program

**UNITED STATES
NUCLEAR REGULATORY COMMISSION
WASHINGTON, DC 20555-0001**

**OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, \$300**

**SPECIAL STANDARD MAIL
POSTAGE AND FEES PAID
USNRC
PERMIT NO. G-67**